TECHNICAL REPORT

PROJECT TECH TOP
STUDY OF LUNAR, PLANETARY AND SOLAR TOPOGRAPHY

By: D.A. Richards, J.D. Gallatin, and R.O. Breault

CAL No. VC-2104-D-2

Prepared for:

National Aeronautics and Space Administration
Electronics Research Center
Space Optics Laboratory
Cambridge, Massachusetts 02139

Final Report
Contract No. NAS 12-19
July 1967

CORNELL AERONAUTICAL LABORATORY, INC.
OF CORNELL UNIVERSITY, BUFFALO, N.Y. 14221

CORNELL AERONAUTICAL LABORATORY, INC.
BUFFALO, NEW YORK 14221

PROJECT TECH TOP
STUDY OF LUNAR, PLANETARY AND SOLAR TOPOGRAPHY

BY:

D.A. RICHARDS
J.D. GALLATIN
R.O. BREAULT

CAL REPORT NO. VC-2104-D-2
FINAL REPORT
CONTRACT NO. NAS 12-19
JULY 1967

PREPARED FOR:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
ELECTRONICS RESEARCH CENTER
SPACE OPTICS LABORATORY
CAMBRIDGE, MASSACHUSETTS 02139

TABLE OF CONTENTS

TABLE OF CONTENTS, Contd.

TABLE OF CONTENTS. Contd.

LIST OF FIGURES

LIST OF FIGURES. Contd.

LIST OF FIGURES. Contd.

LIST OF TABLES

LIST OF TABLES. Contd.

SUMMARY

This is the final report of a program (Project TECH TOP) to assess the state of the art and to make recommendations for future research in topographic information collection systems for planetary and solar investigation. The present report includes a brief review of the first phase effort and a more extensive presentation of the results of the second phase.

The second phase continued analysis of planetary topography to include the initiation of a detailed system analysis, a review and analysis of image restoration processes and their application to space imagery, and some mapping considerations for a stereo imagery system.

Three major aspects of solar topography, theoretical, engineering and experimental, were investigated. The theoretical aspects concerned the concepts, phenomena and problems of major importance to the study of the sun. The engineering aspects concerned solar probe feasibility and survivability, the use of cooperating vehicles, and earth-based solar instrumentation. Two experimental possibilities were examined involving the use of an active backscatter probe of the sun and the use of Fourier transform spectroscopy in space.

Specific recommendations for future research and experimentation are included in the two main sections on planetary and solar topography. Fifteen appendices contain detailed discussions of the various topics covered during the second phase. A bibliography concludes the report.

1. INTRODUCTION

This Final Report describes the work performed under NASA Contract NAS 12-19 (Project TECH TOP) during the period from 1 July 1965 to 15 February 1967. This effort was divided into two phases, the first extending from the beginning of the contract until 15 April 1966, and the second consisting of the remainder. The results of the First Phase have been reported in detail in a formal report*. and are therefore only briefly reviewed in this document.

1.1 Objectives

The objectives of Project TECH TOP were to study available methods, and to develop concepts of new techniques, for acquiring, transmitting and displaying three-dimensional information of lunar, planetary and solar topography. Comparisons of the various techniques and recommendations concerning research and instrumentation required to promote the state of the art were other requirements.

1.2 Tasks Accomplished

The objectives of this program are very broad in scope and consequently all the aspects and disciplines involved could not be scrutinized to the depth desirable. The specific achievements of the entire (both phases) program are summarized here.

During the first phase, the methods of data acquisition that might apply for lunar, planetary and solar surfaces were described. **A** more detailed investigation of one particular method, namely, stereoscopic imaging, included consideration of the requirements, sensors and errors involved. The transmission and display of three-dimensional information were examined in terms of requirements and methods. The applicability of holographic methods was assessed for acquisition, transmission and image processing.

---

* REF. 1 - Campbell, Charles E., "Project TECH TOP - Study of Lunar, Planetary and Solar Topography", First Phase Report, CAL Report No. VC-2104-D-1, August 1966, **(NASA** CR-650).

Other subjects surveyed included the peculiar problems of changing solar topography, measurements to test the general theory of relativity that involve the sun and measurement of stellar parallax from a space vehicle.

During the second phase, analysis continued of planetary topographical information collection requirements to include the initiation of a detailed system analysis, a review and analysis of image restoration processes and their application to space imagery, and some mapping considerations resulting from the first phase investigation.

Solar topography was investigated in greater depth with respect to the theoretical, engineering and experimental questions one encounters in its study. The theoretical aspects concerned the concepts, phenomena and problems of major importance to solar physics. The engineering or practical aspects concerned solar probe feasibility and survivability, the use of co-operating vehicles, and earth-based solar instrumentation. Two experimental possibilities were examined involving the use of an active backscatter probe of the sun and the use of Fourier transform spectroscopy in space.

Specific recommendations for future research and experimentation were generated throughout the program. Numerous references were also compiled.

1.3 Report Organization

This report is divided into four main sections besides the introduction. They are: a review of the First Phase; subjects of concern primarily to planetary topography; subjects related to solar topography; and the appendices.

Section 2 presents a brief review of the First Phase of the program. It includes a summary of the areas that were considered, the recommendations derived therefrom, and the plan for study under the Second Phase.

Section 3 presents those aspects of our investigation which are concerned with the acquisition, transmission and processing of three-dimensional information of a planetary surface. (To avoid continuous repetition of the phrase "planets and satellites" the word planetary will be

used in a generic sense to include all the bodies of the solar system with mappable surfaces such as the planets, their satellites and the asteroids, and to exclude comets and the sun.) The major topics treated here are image restoration and system analysis. The resulting recommendations from this study and analysis conclude the section.

Section 4 summarizes our review of the problems, theories and several major features of solar topography. It also covers some of the instrumental difficulties and experimental possibilities involved in solar research. Those conclusions and recommendations which are derived from these considerations terminate this section.

The final section contains the detailed appendices which support the material of the planetary and solar sections in a more substantial form. Individual topics of research are maintained as separate units to provide easier referencing and to allow readers from the various disciplines represented here to study in greater depth any specific subject of interest without being encumbered by the others. The final appendix is a bibliography.

References in the main body of this report are included as footnotes. Where no references appear, the reader may find in the more detailed appendices more expansive discussions and also some general references.

1.4 Acknowledgements

The authors wish to acknowledge with gratitude the efforts of those who contributed to the First Phase of the Program as well as those of our colleagues at Cornell Aeronautical Laboratory who have aided us with their counsel during the Second Phase. Acknowledgement is also made of the understanding sponsorship afforded us by the personnel of the Space Optics Laboratory, Electronics Research Center, National Aeronautics and Space Administration.

2. REVIEW OF FIRST PHASE

The First Phase of Project TECH TOP has been documented in a formal report (Cornel1 Aeronautical Laboratory, Inc. , Report No. VC-2104-D-1, August 1966). A brief description of its content includes the following: the methods of data acquisition that may apply for lunar, planetary and solar surfaces; a detailed investigation of the stereoscopic imaging method including requirements, sensors and errors involved; the applicability of holographic methods; the transmission and display of three-dimensional information in terms of requirements and methods; a limited survey of solar topography; a review of some experiments to test the general theory of relativity involving the sun; and an analysis of the problems involved in the measurement of stellar parallax from a space vehicle.

The remainder of this section discusses the areas of consideration mentioned above, the recommendations derived from those considerations and the areas that were planned for investigation during the second phase.

2. 1 Areas of Consideration

During the first phase of the program all of the subjects mentioned above were covered, although because of the broad spectrum of subjects included, it was not possible to treat all of them to as great a depth as might have been achieved in an effort of more limited scope.

A brief review of the possibilities of acquisition techniques was presented and is summarized in Table I.

In the area of topographic data acquisition, emphasis was given to stereoscopic (or parallax) methods, which require two images of the surface obtained from two different observation points. In this study consideration was given to: (1) the requirements for obtaining accurate topographical information from stereo imagery (summarized in Table 11), (2) the various sensors which are presently available or future possibilities for obtaining this imagery, and (3) the errors involved in such a collection effort and ensuing data reduction.

Table I
POSSIBLE METHODS OF OBTAINING TOPOGRAPHICAL INFORMATION

1. STEREOSCOPIC IMAGING
    a) PHOTOGRAPHIC
    b) RADAR
2. HOLOGRAPHIC
    a) LASER
    b) MICROWAVE
3. RANGING
    a) LASER
        1) SCANNING
        2) PROFILOMETER
    b) RADAR
        1) ORDINARY
        2) INTERFEROMETER
4. PHOTOMETRIC
    a) KNOWN REFLECTANCE FUNCTION
    b) SHADOWING
5. SURVEY
6. INFERENCE

Table II

REQUIREMENTS TO OBTAIN TOPOGRAPHICAL INFORMATION BY STEREOSCOPY

1. TWO STATIONS FOR TWO IMAGES WITH PARALLAX
   a) CONSECUTIVE
   b) SIMULTANEOUS
2. DISTANCE DETERMINATION
   a) ALTITUDE
   b) BETWEEN STATIONS
   c) SCALE FROM GROUND OBJECT
3. RESOLUTION REQUIREMENT
4. GEOMETRIC FIDELITY PRESERVATION
5. ATTITUDE DETERMINATION
6. DATA REDUCTION

The investigation of holographic techniques under Phase I included an assessment of the direct use of the three-dimensional properties of holographic images to acquire topographic data from space vehicles, an analysis of the problems of transmitting a hologram via a data link, and a preliminary study of how holographic techniques might be employed to improve the quality of degraded images. A detailed analysis was made of the influence of non-linearities in the hologram recording process on image reconstruction.

The application of holographic techniques to the acquisition of three-dimensional data from space vehicles was also investigated with respect to the laser power requirements and the dimensions of the holograms required to achieve adequate depth perception.

The First Phase work in the area of .data transmission consisted primarily of a review of the state of the art with consideration of the general requirements for space applications, bandwidth compression, and laser communication systems.

In the area of displays, both stereoscopic and non-stereoscopic methods of displaying three-dimensional information were reviewed to determine their applicability to space missions. The various methods were compared with respect to resolution, convenience, distortion and their adaptability to real time situations.

A study of stellar parallax measurements was conducted which included consideration of the implications of improving the accuracy of stellar parallax measurements from the astronomical point of view. The engineering problems associated with obtaining improved stellar parallax measurements from space vehicles operating at distances of 50 to 100 astronomical units from the earth were studied.

The effort in the solar area under Phase I was limited to a survey of the literature to determine the specific types of data that are of major importance. The problems of solar topography differ considerably from those of lunar and planetary topography where the surface features are well defined and do not vary in time. Only preliminary consideration was given

to methods for acquiring the desired data. During the study relating to solar topographic measurements, it became apparent that optical observations from space vehicles might be advantageously employed to verify various relativistic and cosmological theories. The significance of performing such measurements was assessed.

2.2 First Phase Recommendations

Conclusions and recommendations which resulted from the foregoing considerations were included in the First Phase report. The conclusions are not reproduced here as they are closely related to the detailed analysis and study presented in that report, and would appear incongruous alone. However, the recommendations which were derived from those efforts are considered more important to the present reader as they require further action. They are restated here so that all suggestions derived from this program appear in the final report.

Therefore, as a result of the work on the First Phase of this program, it has been recommended that:

1) An investigation of the possibilities of improving methods of imaging such as electrography be made for space applications.

2) A laser rangefinder be used for the altimeter requirement in an imagery collection system.

3) A mapping system use a cartographic camera (for position preservation), a reconnaissance camera (for high resolution detail), a star camera (for orientation recording), and an altimeter (for scale determination).

4) The problem of determinig, as a function of time, the orientation of a planet being imaged for cartography be given serious study.

5) Investigation be continued into the problems involved in employing holographic techniques for the transmission of data to define limitations and establish the areas that require further analysis.

6) An experimental and analytical investigation of image enhancement techniques using holograms be performed to establish basic feasibility and to define limitations.

7) Concerning laser communication, work on solving the pointing problem be given highest priority.

8) A complete system analysis be made for measurement of stellar parallax if a space mission for this purpose is planned.

9) The development of telescopes with large focal lengths be explored and studies of systems which would provide greater stabilization than is currently available, be investigated.

10) Specific research be conducted to solve the sensor-metrology problems associated with accurate measurement of the angles between stars.

11) The preservation of the accuracy (position, tone and detail) of photographic imagery in the adverse environment of space, and the non-destructive readout of selenium plates be investigated.

12) The study of solar topography be continued to determine what should be measured and the difficulties involved.

13) The feasibility of measuring, from a satellite, the gravitational deflection of light by the sun and the oblateness of the sun be studied, and the best sensors and optics for performing these measurements be selected.

2.3 Plan for Second Phase

At the termination of the First Phase of Project TECH TOP it was realized that a more limited scope was advisable. Therefore, an abbreviated list of subjects proposed for study during the continuation was presented and received the concurrence of the sponsor. This included solar topography, cooperating vehicles, image restoration and system analysis.

A major portion of the Second Phase was scheduled to be expended on solar topography, The theories of solar physics, the presently available knowledge about the surface phenomena, and the limits of earthbound observations were to be reviewed. An attempt would be made to define the requirements for substantiation of the theories, to examine new techniques and to delineate the problem areas.

The problems of obtaining topographic information from cooperating vehicles was to be examined to include the types of vehicle situations (with respect to each other and the body being investigated), the methods of achieving cooperation and the relative merits of each system defined.

A review of various methods (including holographic) of image restoration would be made with emphasis on their applicability to the degradations anticipated from space collection systems.

Based on the premise that a knowledge of the interactions of sub-systems in an integrated system would more meaningfully pinpoint "weakest links" and help establish priorities of research and development needs, it was planned to initiate a system analysis approach to topographical information collection.

All these topics were examined to varying degrees during the Second Phase.

3. PLANETARY TOPOGRAPHY

3. 1 Introduction

An important aspect of the space program is the exploration of the celestial bodies in our solar system. One of the major tasks of this endeavor is the collection of information about their surfaces. Such knowledge is a prime requisite to landing missions for planetary investigation. It is also of significance to other disciplines such as astronomy and geology. The detailed description of a planetary surface to include its relief and physical features is called its topography. The techniques and considerations necessary to obtain quantitative measures of planetary topography by means of space vehicles is the subject of this section.

Much of the material which has been covered concerning planetary topographical information collection systems was reported in the First Phase report. However, most of the considerations presented there were subsystem oriented. A part of the continuation was devoted to a system analysis approach to the problem. The results of this endeavor comprise Section 3. 2.

Another major portion of the Second Phase effort was concerned with methods of image restoration, i. *e.*, the processing of space-obtained imagery to make extraction of information therefrom easier and/or more accurate. The types and methods of image enhancement are the subject of Section 3. 3.

Several other topics were investigated as problems of subsystems of a planetary mapping system. The primary questions which were considered in this respect concerned the power requirements for a simple laser rangefinder to be used as an altimeter, and the exposure times necessary for star cameras used to obtain orientation. These analyses are summarized in Section 3. 4.

This section is concluded with those recommendations concerned with the various aspects of obtaining topographical information from a planetary surface.

3.2 System Analysis

3.2.1 Concept - A system may be defined as a group of units so combined as to form a whole and to operate in unison. Emphasis is placed on the fact that an overall operational process is under consideration rather than a collection of pieces. If one desires to implement or at least specify an optimal system for a particular task, a systematic procedure should be followed to attain that goal. This is particularly true for space exploration where the need for "crash programs" is not as necessary as some military projects. Such a procedure can be divided into basically eight steps. They are: (1) define the objectives, (2) establish constructs for their achievement, (3) project the subsystems of these constructs, (4) list their constraints, (5) build a mathematical model to interrelate the important parameters, (6) analyze the tradeoffs available within the system, (7) obtain a detailed description by iteration of this process, and (8) achieve an optimally designed system.

Of course, the starting point is to carefully define the mission objectives of the system as given by the ultimate user and refined by the system designer. Selection of improper criteria for optimization means expending effort for the solution of the wrong problem. We may start with tentative objectives which are further refined by feedback from the study itself.

Constructs are then developed which represent possibilities for the solution of the problem. These are hypothesized systems which are candidates for the overall concept to be used.

The process of reducing a given system to subsystems is called projection. The system constructs are projected into the various combinations of subsystems that are conceivable and admissable. A subsystem, in turn, can admit of several possibilities also.

Once a particular subsystem is chosen for detailed analysis, a list of the constraints imposed upon it should be made. These include system requirements which arise from mission specification, physical

constraints which are limitations of the materials and components used, and design constraints which are a result of secondary considerations such as economy or size.

After obtaining a fairly detailed description of the subsystems in terms of the various key parameters involved, the next step is to attempt to combine the descriptions of the individual subsystems into a mathematical model of the total system. Such a representation should relate inputs, outputs, and system parameters.

Then, as the model becomes more sophisticated, one can vary the separate critical parameters to determine their particular effect on overall system performance. Tradeoffs can be specified and decided upon.

Any attempt to thoroughly analyze a system usually proceeds by iteration. Each item is described roughly at the start and then, as the above steps are proceeded through, increasing detail is added; the cycle repeats until the desired definition of the entire system is attained.

Finally, after sufficient analyses of the various alternatives, an optimal system for the original objectives and constraints can be selected.

For complex systems this procedure can become very involved. Not all subsystems are readily describable by mathematical analogues. Sometimes the mathematics becomes too complicated to allow exact solution and approximations are required which compromise the optimality of the result. The relative amount of effort expended on such an analysis will depend a great deal on the importance and cost of the mission.

3.2.2 Considerations - For the system of primary concern to this program, namely, a planetary topographical information collection system, it is possible to delineate many subsystems. Each subsystem will have its own unique considerations of requirements and constraints. Let us take, for example, the system that was looked at more deeply during

the First Phase of this program, namely, stereo imaging. For such a system one can distinguish three main divisions of five subsystems each. The divisions are acquisition, transmission, and processing (see Table III).

In the acquisition phase there are basically five subsystems to be concerned with. They are the subject, the observation station, the sensor, storage and readout, The subject (planetary surface) is outside our design control and primarily contributes constraints to the system which are inflexible. The observation station or platform (space vehicle) requires stability, attitude and altitude determinations and will most probably be limited by size and weight. Sensor considerations include limitations on the collector, sensitive element, filter and integration time. Some form of storage can be considered as always necessary for imagery, be it just the integration time of the sensitive element or a longer time necessary because of transmission problems. The readout subsystem offers many alternatives including analog vs digital, multi-channel vs scanning, as well as the different types of scanning possible. These must all be compared to determine which is optimum for a particular system. Table IV summarizes the major considerations of the acquisition division subsystems.

Once a signal has been produced, we enter the transmission division of subsystems. This consists of the encoder, the transmitter, the channel, the receiver and the decoder. The encoder subsystem includes bandwidth compression and modulation choices. The transmitter requires determination of power, carrier frequency, and antenna limitations. The channel exhibits natural constraints of distance and noise. The receiver subsystem includes the receiving antenna size requirement and receiver sensitivity. The decoder consists of the detection and demodulation phases which lead to reconstruction of the original signal.

After obtaining the reconstructed imagery, there are several operations that can be performed which are grouped under the title of processing. Five subsystems in this division can be specified. They are restoration, extraction, evaluation, feedback and use. Restoration considerations include tone or grey-scale, position and detail in the imagery. Extraction requires the display or data reduction machines and methods to

Table III

DIVISION AND SUBSYSTEMS OF A SYSTEM

PLANETARY TOPOGRAPHY SYSTEM USING STEREO IMAGERY		
ACQUISITION	TRANSMISSION	PROCESSING
PLANET	ENCODING	RESTORATION
PLATFORM	TRANSMITTER	EXTRACTION
SENSOR	CHANNEL	EVALUATION
STORAGE	RECEIVER	FEEDBACK
READOUT	RECONSTRUCTION	USE

Table IV

THE SUBSYSTEMS OF THE ACQUISITION DIVISION OF A STEREO IMAGERY SYSTEM

1.	PLANET	a) LIGHTING b) ATMOSPHERES c) ORBIT d) ALBEDO
2.	PLATFORM	a) STABILITY b) AIMING DETERMINATION c) ALTITUDE DETERMINATION d) ATMOSPHERIC ENTRY
3.	SENSOR	a) COLLECTOR b) SENSITIVE ELEMENT c) FILTERS d) INTEGRATION TIME
4.	STORAGE	a) DEVELOPMENT b) MAGNETIC c) ON-BOARD PROCESSING d) STORAGE TUBES
5.	READOUT	a) ANALOG vs. DIGITAL b) SCANNING c) MULTI-CHANNEL d) RETURN

be analyzed to determine the limitations they impose. Evaluation is that portion of the information processing which enables us to say how well we are meeting the original objectives. Feedback is necessary if the system is to be used more than once. It supplies the inputs to adjust, modify or replace subsystems for subsequent operations. The final process the imagery is subjected to is the use for which all the rest of the effort was utlimately intended. It is this subsystem which requires the greatest care in analysis as it in turn supplies the objective requirements.

During the Second Phase of this program primary concern in the area of detailed system considerations, has been with the problems attributed to planetary peculiarities (see Table V), and to the combination of elements (optics-sensor-readout) at the front-end of an imagery collection system. Some of the results of this analysis are contained in Appendix P-1.

3.2.3 Model - The considerations outlined for the subsystems should carefully define the critical parameters and attempt to relate them mathematically. Then the next step is to combine these subsystem descriptions into an overall mathematical model of the total system. Such a representation should relate inputs, outputs and system parameters. By programming the model for simulation on a digital computer, the detailed investigation of system performance under varied combinations of conditions can be carefully studied. Such simulative techniques represent a powerful and economical method of achieving optimal system performance, and establishing design compromises among components prior to the initiation of costly hardware programs. The ultimate problem of such an analysis is to determine the actual transfer function (spectral and spatial) of each functional unit in the system and to appropriately combine these into a single equivalent transfer function. Figure 1 illustrates a simplified combination of component (subsystem) responses to produce an overall system transfer function for spatial frequencies. The synthesis here is merely the multiplication of the individual curves to produce the system response. (Several more transfer functions are involved in the actual case of transfer from source to user.)

Table V

KEY PARAMETERS OF THE PLANETS AND THE MOON

NAME	DISTANCE FROM SUN (A. U.)	MASS (Me)	RADIUS (KM)	LIGHT (Lp/Le)	ALBEDO	ATMOSPHERE
MERCURY	0.387	0.0543	2480	6.7:1	0.055	NEGLIGIBLE
VENUS	0.723	0.81 36	6200	1.9:1	0.80	DENSE
EARTH	1.00	1.0000	6371	1:1	0.35	TRANSPARENT
MARS	1.524	0.1080	3310	1:2.3	0.15	TRANSPARENT
JUPITER	5.203	318.35	70,000	1:27	0.44	DENSE
SATURN	9.539	95.22	57,500	1:91	0.42	DENSE
URANUS	19.19	14.58	25,500	1:370	0.45	DENSE
NEPTUNE	30.07	17.26	24,800	1:900	0.52	DENSE
PLUTO	39.46	0.93	3,000	1:1560	0.16	UNKNOWN
MOON	1.00	0.0123	1738	1:1	0.07	NEGLIGIBLE

1 ASTRONOMICAL UNIT (A.U.) = 1.496 x 10^8 km.

1 EARTH MASS (M_e) = 5.976 x 10^{24} kg.

L_p/L_e IS THE RATIO OF ILLUMINATION AT A PLANET TO THAT AT THE EARTH.

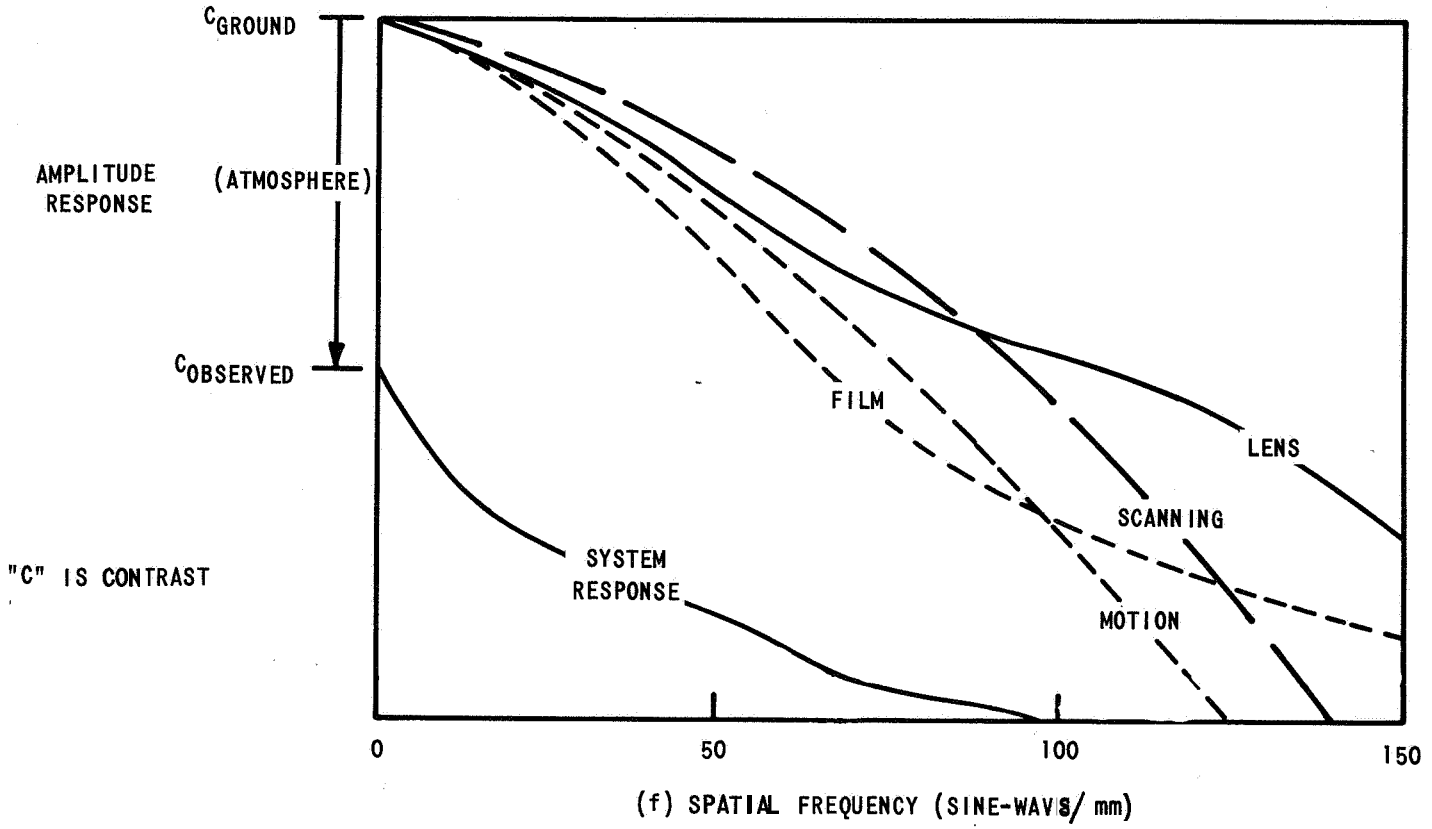


Figure 1 REPRESENTATIVE SYNTHESIS OF SYSTEM RESPONSE FUNCTION

By programming the various possibilities and the combinative process, one can generate the system response of a great multitude of combinations so that tradeoffs for cost-effectivness can be quickly and adequately analyzed. A sample* of such variations for lens-film combinations is presented in Figure 2. This illustrates the response changes of a KA-30 reconnaissance camera equipped with a six-inch Xenotar lens using Plus-X Aerecon emulsion when subjected to different motion conditions (stabilization and image motion compensation error) for two aperture settings.

The procedure described above requires not only accurate mathematical representations of subsystem responses but also accurate data inputs concerning the various components available. If no analytical relationships can be found which describe a particular component's transfer function, information derived from experiment and testing presented in graphical form is used (as in Figures 1 and 2). The compilation of sufficient data of this type is a formidable, but not impossible, task. The most difficult aspect in this respect is the representation of subsystems not readily analyzed mathematically, such as planetary atmospheres and man (who enters the chain as a stereoplotter operator when the use of the imagery is mapping). However, it is our firm conviction that this powerful technique is the logical approach to indicate the direction of design choices and modifications for optimization, and also to point out fruitful areas for research and development effort.

3.3 Image Restoration

3.3.1 Introduction - This section discusses the ubject of image restoration. It includes a review of the basic methods available and some more particular application problems. Several appendices (P-2, P-3, and P-4) treat these areas in greater depth so they are merely summarized here.

---

* REF. 2 - Kaufman, Sol, "A Mathematical Model for Analysis of Reconnaissance Systems with Specific Application to Scientific Space Reconnaissance", CAL Internal Report, 18 November 1959.

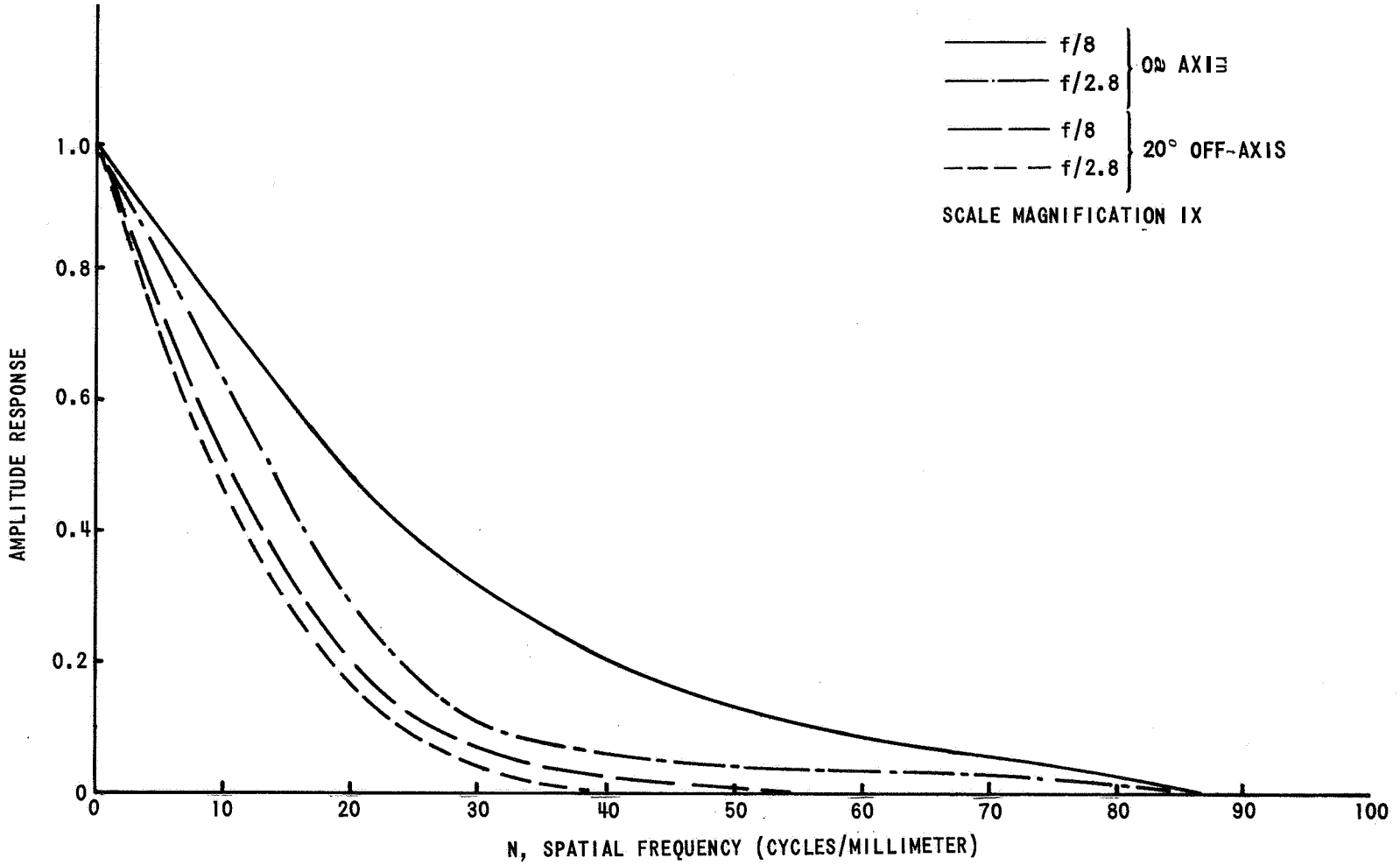


Figure 2

LENS-FILM RESPONSE FUNCTIONS

KA-30 6" XENOTAR LENS PLUS-X AERECON EMULSION

Image restoration (also called image processing or image enhancement) is that operation performed on imagery that has been subjected to some distortion or degradation, which attempts to correct or compensate for these detrimental effects by a subsequent process. However, it must be remembered that* because of limitations on resolution and because of the presence of noise, any image contains only a finite amount of information. Image degradation always reduces the information content of an image. In the presence of degradation, the user is incapable of efficient extraction of that information which remains in the image. Processing attempts to re-arrange this remaining information for more efficient extraction. Ultimate processing would allow extraction of all of the remaining information in a degraded image.

There are basically three characteristics of imagery which can be subjects of restoration. They are tone, position and detail. Methods of restoring them are summarized in the following subsections. A more detailed analysis is given in Appendix P-2.

3.3.2 Tone (Grey Scale) - Restoration of tone (or grey-scale) means processing to obtain an image in which the reflectance (or some scalar) is proportional to a suitably defined reflectance of the object. There are several methods for accomplishing this; notably, masking techniques, the use of special film (e.g., gravure film), electrical filtering (for one-dimensional sampling) and computer processing (for two-dimensional sampling). When necessary, tone restoration is usually performed as a preliminary processing operation prior to digital processing for other purposes. Thus, digital techniques may be preferred.

3.3.3 Position - Restoration of position (elimination of distortion) means the process of warping the image so that objects of the original scene are located according to a simple mathematical projection.

---

* Adapted from REF. 3 - Harris, J. L., "Image Evaluation and Restoration", JOSA, 56, No. 5, May 1966, p. 569.

The available methods to do this are: (1) The Porro-Koppe Method which produces a copy using a lens whose distortion is the same as the distortion to be corrected. This method is inapplicable unless a standard lens is used or the taking lens is recovered. (2) Specially figured plates are sometimes used in projection prints. They reduce image detail. (3) Mechanical methods in which different parts of the format are printed by changing the position of a projection lens. These methods either are precise at only a few points or reduce detail. (4) If the image is represented by discrete samples, it may be restored by interpolation. By using the diffraction function and interpolation with a sufficient number of points, only negligible effects on image detail occur. (5) Methods exist to apply correction to the plotter as the map is drawn. If properly implemented, these methods are precise and do not destroy detail. (6) It is possible to use distortion data to correct the map after it has been drawn although such methods can be less precise than those mentioned above.

3.3.4 Detail - Detail restoration is the enhancement of high resolution information available in an image. Several methods are available to process images to achieve this. They operate as follows:

1) Incoherent cross-correlation operates in the space domain. Correlators are relatively simple to build and maintain. They suppress noise and unwanted structure in an image to bring about enhancement of the desired high spatial frequencies. They cannot be used to create point spreads with negative parts without special difficulties.

2) Coherent optical spatial filtering and holographic image restoration operate in the spatial frequency domain. Spatial filtering with the filter on a carrier is the same as holographic restoration. If the filter function is limited to real values, no carrier is necessary. These methods can be used to build very sharp cut-off low pass filters. Therefore, much of the grain noise of photographs which occur in the high spatial frequencies can be eliminated. No method is apparent at the present time for building spatial filters for restoring imagery that has been degraded differently for different points in the format.

3) If the image is represented as a field of samples, it may be restored by numerical filtering. Such processing can be very versatile. However, it should be preceded, and perhaps followed by other filtering to prevent folding of noise* and to restore a continuous image.

There are other methods which are less versatile than these and which may be considered as combinations and special cases of these. Because each has its own limitations, more than one may be necessary. Also, it may be possible to convert an existing device into a processor. Perfect restoration of image detail is not possible for images degraded by noise (the usual case).

3.3.5 Other Considerations - When images are to be used for stereoplotting, for topographical map production, we can consider two cases, namely, human operators and automata. For automatic stereo-compilation, restoration schemes may be chosen to minimize mean square relief error. This analysis is presented in Appendix P-3. When a stereo-plotter operator is involved, the choice of restoration scheme must be based on psychological results. Experimental research is required to determine what to do with degraded images which are to be used in instruments involving human operators.

Several examples of processing space imagery for enhancement have been generated at Cornell Aeronautical Laboratory as a result of an internal research program. They are included in Appendix P-4 of this report with a brief description as illustrations of the processes mentioned above.

3.4 Mapping Considerations

The requirements for a knowledge of orientation and altitude for a stereo imagery mapping system can be satisfied by a star camera and a laser rangefinder, respectively. During the Second Phase of this program

---

* REF. 4 - Blackman, R. B. and J. W. Tukey, "The Measurement of Power Spectra", Dover Publications, Inc., New York, 1958, p. 31f.

a study was made of the star camera exposure times necessary and their attendant problems. A deeper analysis of the power requirement for a simple laser rangefinder system was also conducted. These analyses are contained in Appendices P-5 and P-6, respectively. They are summarized below.

3.4.1 Exposure Times for Star Cameras - For planetary mapping from a space platform, a star camera was suggested to determine the attitude of the platform (and cameras) for triangulation purposes. It, therefore, became necessary to consider the exposure times required for such cameras. Two systems were considered, namely, diffraction-limited and aberration-limited. The reasons for considering diffraction-limited systems are, first, it is desirable to use a sufficiently slow photographic film to eliminate radiation shielding requirements and, second, it would then be advantageous to use a diffraction-limited lens to concentrate as much energy as possible onto a small area on the film. The motivation for considering aberration-limited systems is the exposure problem which occurs in the other type system.

The analysis included a derivation of the relationship between stellar magnitude and illumination at the entrance pupil of the camera. The equation for this is

$$E_A = (2.67 \times 10^{-6}) \times 10^{-0.4m} \qquad (1)$$

where E_A is the illuminance in lumens/meter^2 and m is the apparent magnitude of the star.

Expression (1) was used to find the required exposure time for a diffraction-limited system which is

$$t = \frac{X \lambda^2 F^2 \cdot 10^{0.4m}}{252 \varepsilon D^4} \qquad (2)$$

where t is the exposure time, X is the required exposure, λ is the wavelength, F is the focal length, m is the apparent magnitude, ε is the efficiency of the optical system, and D is the diameter of the entrance pupil.

These results were then applied to a particular system in a numerical example for a diffraction-limited type. Such an investigation revealed difficulties with image motion compensation, stabilization and lens design.

After the above analysis, a particular aberration-limited system was considered. The required exposure time for this system was found and the time and other characteristics were compared to the diffraction-limited system.

It was concluded that both systems have serious problems. The diffraction-limited system would be required to be very large, to use image motion compensation and to need a high degree of stabilization. The aberration-limited system by using a faster film becomes more susceptible to space-radiation fogging and consequently would require more shielding, a disadvantage in weight.

It is recommended that further investigation be performed: (1) to find the optimum field angle and implied required stellar magnitude; and (2) to try other systems between the extremes analyzed. Although some improvement may result, it is believed that obtaining adequate exposure will result in a large and heavy camera or magazine.

3 2 a Rangefinder Considerations - During First Phase of Project TECH TOP some preliminary analysis was performed concerning the power requirements for laser rangefinders. However, the calculations made then utilized values of the parameters which resulted in the background being negligible as compared to the signal. The question arises, what reduction in the power requirement is possible so that, although the noise is no longer negligible, the false alarm rate is tolerable for the intended mission?

Several complications arise in the analysis which attempts to answer this question. These requirements should be derivable from system constraints and a maximum failure probability specified by the system designers. However, the analyses presented in the literature rely on an approximation which provides an estimate on the low side with no true indication of how underdesigned the system would be. (The original mentioned above was overdesigned.) An approach has been settled upon whereby another approximation, which provides a high estimate of power necessary, and in conjunction with the other, brackets the more rigorous answer sought. A more detailed mathematical exposition of the problem and the approximations is presented in Appendix P-6. A better evaluation of the validity of power requirement calculations can now be made. An example indicated that the literature accepted approximation is low by about 10%.

3.5 Recommendations

This section presents a compilation of the recommendations which resulted from our investigation of the various aspects of obtaining topographical information from a planetary surface on the Second Phase of this program. They are presented in the order of their applicability to the flow of information from source to user as presented in Section 3.2.

They are further classified as to time, period of applicability and type of action recommended. Time period is broken into immediate (I), near (N), and far (F)future. By immediate is meant action which could or should be initiated at the present time. Near future includes the next several years and far future is the seventies and beyond. Action required is divided into experimental (E), study or theoretical (T), and research and development (R).

Total System:

1. Complete a Detailed Analysis of a Total System for a Planetary Mapping Mission (I, T) - The analysis of individual subsystems for such an effort does not give the overview necessary to optimize each element in combination with the others. The chain which starts at the

illumination on the subject planetary surface and is completed with the output of topographical maps can only be as strong as its weakest link. Components should be matched to the extent that high performance achieved in an element early in the system at a great cost should not be compromised by a less than adequate subsequent element. For example, the Mariner IV probe and the Tiros satellite were limited in ground resolution, the former by its readout scanning system and the latter by its orbital motion. A few more scan lines on Mariner IV or an image motion compensation system on Tiros (both within the state-of-the-art) would have achieved greater ground resolution. The Lunar Orbiter imagery would have been more suitable for mapping purposes had the designers included a reseau grid like that utilized on the Rangers for geometric fidelity reference.

Planets:

2. Analyze Planetary Motions to Determine Optimum Times for Mapping to Avoid Major Areas of Total Darkness (I,T) - Because of the inclination of the planetary rotational axes to the plane of their orbits about the sun, the phenomenon known as arctic night on earth also occurs on the other planets. The area of total darkness is a function of the inclination, and the time of total darkness is dependent upon the planet's period of revolution. Although this is not a critical problem for the terrestrial planets, it could be one for the major planets. For example, Uranus has an inclination of 98" and a period of 84 years. The next time no major area of darkness occurs in the year 2008. Therefore, mission planning for such tasks must begin with sufficient lead time (minimum energy flight time to Uranus is 16.1 years) to achieve optimum conditions for the data collection.

3. Investigate the Problems of Mapping Asteroids (I,T) - The minor planets or asteroids may provide excellent platforms for extended duration experiments or act as half-way stations in exploring the outer planets. However, because of their relatively small size it may prove extremely difficult to establish satisfactory orbits about them to observe their surfaces. Their masses are not known so orbits cannot be predicted in advance.

Atmospheres:

4. Support Advanced Feasibility Studies for Aerodynamical Vehicles to be Used in Foreign Atmospheres (F,T) - There exists the possibility that cloudy-atmosphere planets might have transparent lower layers similar to the earth. To map 'such planets with optical systems would require vehicles similar to our airplanes. The different atmospheric density and planetary gravity would most probably require special designs.

5. Examine the Possibilities of Learning About the Planetary Topography of those Planets with Non-Transparent Atmospheres by Obtaining Detailed Imagery of the Cloud Formations (N, T) - From an analysis of the Tiros Program and a study of the correlations of topography with major weather systems, it may be possible to infer gross features of the surface from patterns observed in a planet's cloudy atmosphere.

Sensor:

6. Investigate the Use of Radargrammetry for Mapping of Planets with Optically Dense Atmospheres (I,T) - Radargrammetry is the microwave equivalent of photogrammetry. It uses two radar-obtained images to derive topography.

7. Consider a Multi-Sensor Package to Obtain Information in Addition to Geometrical Topography (N, T) - It may not be sufficient to merely know that an area is relatively flat in landing site reconnaissance. Whether it is solid or liquid may be of concern also. Whether the surface observed is the top of dense vegetation or indeed the solid planetary surface is another question which might be answered by an appropriate combination of sensors to assist the topographical imaging devices.

Storage:

8. Determine what Processing Should be Accomplished Prior to Transmission (N, T) - It may be possible to eliminate much information from planetary imagery which is not necessary for the mapping accuracy required prior to transmission with a resultant saving in bandwidth or

transmission time. Also, some form of enhancement before transmission may prove less susceptible to degradation from the subsequent elements.

Readout:

9. Analyze Various Methods of Readinp-Out an Image by Scanning to Determine which (if any) is Less Susceptible to Distortion and/or More Amenable to Restoration (I,T) - Some scanning systems introduce more degradation of imagery than others. The profile of the readout spot is important in terms of flatness of field and also sampling theory. The reconstruction raster's deviation from the readout raster is just another source of degradation of system fidelity. Although much analysis has been done on scanning problems many questions remain unanswered and more work is required.

Transmission:

10. Explore Use of Relay Stations for Distant and Difficult Communications (I,T) - Cooperating vehicles can provide links for distant (edge of the solar system) and difficult (other side of the sun) communications. Study should be performed to determine the problems in establishing and maintaining them.

Restoration:

11. Establish a Systematic Approach to Image Enhancement Applications (I, T & E) - It is recommended that the following four tasks be considered in proposed applications of image enhancement. They are: (1) determine a criterion for judging the quality of the images (depends on intended use), (2) determine the optimum enhancement filters and the theoretical limit of possible enhancement, (3) establish comparisons among the practical limitations of various implementations of enhancement, (4) evaluate the utilization of these techniques in terms of extending original system capabilities. Such an approach requires more analysis and experiment. Some more specific recommendations for image enhancement follow.

12. Perform Laboratory Experiments to Test Current Theories on Resolution Beyond the Rayleigh Limit (I, E) - Recent papers by C. W. Barnes* have presented theoretical analyses indicating the possibility of restoring an object resolution beyond the diffraction limit when the illumination in the object space is confined to a finite region. A system to implement this scheme would be very sensitive to noise. However, experiments should be performed to test this approach since it might be of value for investigation of the dark side of planets.

13. Investigate the possibility of Removing Non-Essential Information from Space Imagery by Spatial Filtering (N, E & T) - If the purpose of examining high resolution imagery is to find potential landing sites, it might be of value to be able to remove from this imagery microstructure (e. g., craters less than two feet in diameter) which makes such an evaluation more difficult. It is conceivable that spatial filtering could be utilized to perform such processing. Spatial filtering can also be used to remove the structure of a raster scan.

14. Investigate the Possibilities of Image Restoration Point-by-Point (N, E & T) - The degradations received by various points in an image are not the same in real systems. This is especially true for images degraded by non-uniform motion. Any restoration processing should take into account the peculiar degradations each point receives.

Extraction:

15. Perform Psychological Experiments with Stereo-Plotter Operators to Determine Image Requirements and Accuracies Obtainable (I,E) - In the production of a topographical map, the ability and experience of the stereoplotter operator enters the system. It is not presently known what image degradations are intolerable or what restorations are of value to the plotter user. A more accurate description (in mathematical terms) of the errors introduced by stereoplotter operators would also be of value to a total system analysis. Such answers require psychological experiments with the people of appropriate training.

* REF. 5 For example, "Object Restoration in a Diffraction-Limited Imaging System", J. Opt. Soc. Am., Vol. 56, No. 5, May 1966, pp. 575-578.

4. SOLAR TOPOGRAPHY

4. 1 Introduction

In a study program devoted to techniques of topographical information collection, solar topography must be treated separately. Because the sun is a gaseous dynamic body, the concept of surface has a less definite meaning than it does for a relatively stable planetary surface. Nevertheless, the practical importance of the sun to the sustenance of life on this planet, and the many astronomical and astrophysical aspects of its stellar structure strengthen the motivation to study and understand its observable regions and the physical features which have been discovered therein. The results of our investigations into these aspects of solar topography including some practical considerations for solar research are the topics of this section.

The description of the phenomena and theories of the sun with their attendant problems comprises the major portion of the literature on the sun. A summary of our review of these problems and theories is presented in Section 4.2. There are many considerations involved in observing the sun from a space probe or from the earth. Both of these cases were studied during the Second Phase, and the use of cooperating vehicles in space applications was also considered. Detailed accounts of these investigations are contained in Appendix s-5, s-6, and S-7. They are summarized in Section 4. 3.

Two experimental possibilities with application to solar research were also examined. The use of a laser probe to study the solar atmosphere by backscatter analysis was evaluated. The value of Fourier transform spectroscopy in solar spectral analysis and the problems of implementing such instrumentation for use in space were assessed. The results of these studies are reported in Section 4. 4.

This section is concluded with those recommendations concerning solar topography which were derived from the above mentioned investigations.

4.2 Theoretical Aspects

4.2.1 Concepts - The sun is a dwarf star of spectral type G, which is in the middle of the range of spectral types of stars. Its absolute luminosity is neither very high nor very low, nor has it any marked stellar peculiarities. The sun is a very ordinary star. It can be considered to be composed of two concentric regions, its atmosphere and its interior. Nothing is known from direct observation about the interior. Electromagnetic radiations are only obtained directly from the atmosphere.

There are three major regions which comprise the atmosphere of the sun. In the order of their increasing distance from the interior they are the photosphere, the chromosphere and the corona. The photosphere is the layer from which the visible solar radiation is directly (without further re-emission) transmitted. It is a thin layer (only about 400 km thick) situated at a central distance of 6.96×10^5 km (called the solar radius). It should be remarked that if the sun is observed in other wavelengths than those of ordinary light its apparent radius may be different. In the photosphere the density falls very rapidly with increasing distance from the center of the sun so that the solar limb (the edge of the visible disk of the sun as seen by an observer) is very sharp. The radiation from inside the limb forms a continuous spectrum with the Fraunhofer absorption lines superimposed.

Above the photosphere comes the chromosphere. When seen at the limb it gives a spectrum of emission lines ("flash spectrum") the intensity of which is much lower than that of the photosphere. The chromosphere has an irregular, rapidly changing, structure and is considered by some authors as consisting of a multitude of small jets or prominences. The height of the chromosphere is about 10,000 km. The corona is situated outside the chromosphere. It has a continuous spectrum due to scattered photospheric light and, superimposed upon it, some emission lines. These derive from very highly ionized atoms. The high ionization indicates that the coronal temperature is about 10^6 °K. The corona can be traced out to about 10 times the solar radius.

4.2.2 Phenomena - There is a multitude of phenomena which are identifiable in the various regions of the solar atmosphere. These different forms of activity can be classified in two manners. First, they can be distinguished according to which layer of the solar atmosphere they are found in. Second, they can be grouped as to whether they are temporary, anomalous features of continuous, more-or-less steady state features. In the latter category, one can place the differential rotation, granulation, and spicules. Besides these, nine other variable phenomena will be briefly described here as the observable features of solar topography.

Observations of the motion of various features (primarily sunspots) of the photosphere have indicated that different latitudes rotate with different angular velocities (faster at the equator than at the poles). These observations of differential rotation are confirmed (although not precisely) by Doppler measurements at the limb. Such measurements in the chromosphere indicate that the angular velocity increases with height above the photosphere also. A discussion of this peculiarity is presented in detail in Appendix S-1.

The other continuously present aspect of the photosphere is granulation. The normal photosphere is not uniformly bright but shows a cell-like structure of irregularly shaped bright areas (on the order of hundreds of kilometers in diameter) surrounded by darker boundaries. Many difficulties remain in any description of granulation which purports to be complete. An examination of the theories and observations to date is contained in Appendix S-2.

Probably the best known, although not best understood, features of the photosphere are the sunspots. They are the most conspicuous signs of solar activity and have been known since Galileo's time. The number of spots observable at any given time varies in an eleven year cycle. The spots appear on the sun in two zones at the beginning of a cycle, at about 35°N and 35°S latitude. Over the course of the cycle, the spot groups move further towards the equator (±8° latitude at the end). All spots show magnetic fields (up to 4000 gauss or more) with the maximum field intensity,

lying at the center. At present, no satisfactory theory in all respects concerning spots, spot groups or the cycle exists. Stereo pairs of the sun taken from the earth (parallax obtained by rotation of the sun) show the spots to appear as depressions.* No extensive investigation of sunspots was performed during this program, however an excellent treatment of the subject is available.**

When near the solar limb, sunspots are seen to be surrounded by irregular bright elevated areas called faculae. Elsewhere they become less conspicuous because of the brighter background. All spot groups are surrounded by facular regions, although there are also many such regions without spots. The origin of faculae is not at all well understood, but they are undoubtedly formed as a result of magnetic disturbances. These cause some regions of the photosphere to be heated in excess of the average temperature and to rise for several miles above the surrounding level. Polar faculae and chromospheric faculae have also been observed. It has been suggested that the polar faculae have some connection with the polar rays of the corona and perhaps the general solar magnetic field. The chromospheric faculae are observed in the metal line emissions and appear to be the extension of the photospheric faculae. The chromosphere, when seen in profile at the limb during an eclipse, appears to consist of a roughly level region about 4000 kilometers high surmounted by fairly closely packed spicules - small jetlike protrusions with lifetimes of the order of several minutes and upward velocities of about twenty kilometers per second. Spicules are a normal phenomenon of the sun since they occur in undisturbed regions and at all times. The appearance of the chromosphere in the light of H_α that arises from this structure has been compared to a prairie fire.

---

* REF. 5 - Menzel, D. H. and G. Moreton, "Spectroscopy from Space Satellites", Space Age Astronomy, Academic Press, Inc., New York, 1963, p. 412.

** REF. 6 - Bray, R. J. and R.E. Loughhead, "Sunspots", John Wiley and Sons, Inc., New York, 1965.

Plapes are chromospheric phenomena usually viewed in the light of the calcium lines. They are also called chromospheric faculae. Solar flares are probably the most important solar feature as far as man is concerned. These chromospheric eruptions can spew forth high energy particles which may affect our weather, our communications and our space exploration. The basic concepts and description of flares are presented in Appendix S-3 along with some ideas to be investigated to promote prediction capability.

Prominences are huge projections of matter extending from the chromosphere into the corona for thousands of kilometers. There are basically three types, namely, quiescent, active-region and surge. A quiescent prominence develops slowly and may last for several months. Its association with other phenomena is not obvious. Active-region prominences are those that are formed in and remain close to sunspot regions. Their temperatures are higher than the quiescent types. They often appear as enormous loops forming arches along what appear to be magnetic lines. Surge prominences form very rapidly by sending a large flux of luminous materials far into the corona in a short time before falling back or fading away.

Thin or blade-shaped prominences, when they appear dark against the solar disk in monochromatic photographs, are sometimes called filaments. This separate designation survives from the time when it was not realized that filaments and prominences are simply two aspects of the same phenomenon.

Coronal condensations are regions of the corona lying above active areas of the photosphere and chromosphere which generally exhibit increased emission intensity and broadening of the spectral lines arising from highly ionized atoms. The observations suggest that the temperatures and densities in these regions are higher than the surrounding corona. The hot spots are called permanent condensations. There are also sporadic condensations corresponding to cool spots.

Streamers (also called fans, arches and rays) are generally any long extension of the corona, broad or narrow, with or without fine structure. A fan is a big streamer, wide at the base and narrow at greater distance. Rays are narrow streamers that occur in a variety of forms.

The above-mentioned phenomena are the primary features of solar topography. More detailed descriptions of these phenomena are available in many books on the sun.* The common aspect of all treatments, however, is the lack of a firm description (theory conforming to observation) more or less for all of them. A good deal of work is still necessary in both observations (in greater numbers and greater accuracy) and theoretical analysis. See Section 4. 5 for more specific recommendations.

4.2. 3 Problems - There are many major unsolved problems in the realm of solar physics. A brief description of some of these subjects of research is given below.

The composition of the sun by elemental abundance is not accurately known. Analysis of solar spectra for an extended period has resulted in identification of approximately seventy of the known elements. The individual abundances vary from one investigator to another. It is also thought that all the elements exist there but have not been seen because of low abundances and lack of sensitivity of the instrumentation used. The atmosphere also limits such observations. We will see the importance of these determinations in Section 4. 2.4.

Although hydrogen and helium are the two most abundant elements in the sun, the ratio of hydrogen-to-helium is not well known. This is an important quantity in the description of stars and in the construction of theoretical models of stellar structure. Only rough estimates (~6:1) exist

* See Appendix B-1. Bibliography.

at the present time. The spectral line analysis for helium is plagued with many difficulties and uncertainties. The necessary experimental parameters (e. g., line wing shape) must be improved.

Tens of thousands of lines have been observed in the electromagnetic spectrum of the sun. Approximately thirty percent of these remain unidentified as to the transitions which produce them. Much of this spectral problem can be reduced with the expansion of experimental programs in transition cataloguing.

Not much had been done until the advent of the space program about the X-ray and ultraviolet radiations of the sun, because of atmospheric absorption. However, data are rapidly mounting as the result of recent rocketry programs to obtain spectra in these regions. Nevertheless, at least half of the lines so obtained have not as yet been identified from the multiplet tables. The infrared region does not suffer from as severe an atmospheric limitation. However, molecular (CO_2, H_2O) absorptions render many quantitative studies in the solar IR less firm. Both regions (above and below the visible) require more exoatmospheric work.

The measurement of a general magnetic field for the sun has not produced conclusive results. It was originally estimated at about 60 gauss.* This was later reduced to the order of one gauss. Presently a theory is put forth which suggests a variable general magnetic field which switches polarity from north to south. The magnetic fields associated with the phenomena discussed in Section 4.2.2 also require more accurate and localized measurements.

The formation and behavior of sunspots has been a puzzle since their earliest observations. They are generally believed to be intricately and intrinsically connected with localized magnetic fields. Precisely what the mechanisms are that produce the structure witnessed in observations has not been established. Both increased optical resolution

---

* The earth's general magnetic field is one-half gauss for comparison.

and greater fine structure resolution of the magnetic fields are needed. Theoretical and experimental magneto-hydrodynamics must also be advanced to arrive at an acceptable description of sunspot morphology.

The structure observed in the corona (condensations and streamers), including the asymmetries which occur during the solar cycle, are not understood. Also, the energy transfer mechanism which supports the high temperatures (~1, 000, 000°K) of the corona outside the relatively low temperature (~5000°K) photosphere is a mystery that has generated many speculative theories. No choice can be made among the several explanations offered in either case until more reliable data and analysis are obtained on the magnetic fields, and the emission lines and continua of the highly ionized atoms which appear to exist there.

The differential rotation which is in evidence on the sun has two major unanswered questions: what is the exact value of the angular velocity as a function of latitude and what is the origin of the equatorial acceleration. The measurements of angular velocity differ for various methods of determination and the theories offered to explain its origin draw little support beyond their originators. More precise measurements from space are needed.

From evidence recently gathered, solar granulation presents the appearance of the cellular pattern associated with non-stationary convection. However, the Rayleigh number* appropriate to the bottom of the photosphere exceeds the critical value by several orders of magnitude which means this convection should be an entirely random phenomenon. Much additional information is needed before the problem of granulation can be resolved. Such a task will necessarily require a high degree of stability in the directional control of the informational collection instruments used.

---

* REF. 7 - Aller, L. H., "The Atmosphere of the Sun and Stars", Astrophysics, Ronald Press Company, New York, 1963, p. 451.

Because of their great importance to man, the onset and mechanisms of solar flares has become the most heavily studied problem of solar physics. A prediction capability for super flares which emit dangerous solar cosmic rays would be of great value to the manned space program. Much more effort in observations and analysis is necessary to obtain this capability.

All these problems and others not mentioned here have basically similar elements. Available data are scattered and scarce. No one theory is accepted by a majority of the investigators as the correct explanation. They all require more observations, preferably outside the atmosphere, and a good deal more analysis (including any experimental work to support the analysis). They also are in need of better theoretical descriptions.

4.2. 4 Modeling - The solar atmosphere has two major aspects of its topography which are amenable to description. One treats the distinguishable features which we have discussed in Section 4.2.2. The other attempts to describe the values of those physical parameters which can be given in the form of a profile in height (or depth), such as temperature, pressure, etc. for all three regions of the atmosphere from which observable emanations radiate. Such descriptions are called models. We are here concerned with solar atmospheric modeling.

Many investigators have constructed such models using various assumptions, data inputs and theories. The requirements include a value for the surface gravity, a knowledge of the relative abundances of the elements and a thermodynamic theory of the radiative process involved. The combinations and permutations of the various building blocks result in the situation that there are as many different models as there are model builders. A review of the procedures and ambiguities involved is presented in Appendix S-4. Recommendations for the alleviation of this situation are given in Section 4. 5.

4. 3 Engineering Aspects

There are many complex engineering considerations to be evaluated before a study of the sun can be practically implemented. To answer the question, "How close can we get to the sun?", rnany requirements and limitations must be scrutinized. Such an analysis for a solar probe is presented in Appendix S-5. **A** summary of that exposition is contained in Section 4. **3.** 1.

The use of cooperating vehicles was one of the modes of stereo acquisition to be considered during the Second Phase of Project TECH TOP. Since the most promising application for such a system was found to be in the area of solar topography, Section 4. **3.2** summarizes the results of that study. The expanded considerations are presented in Appendix *S-6*.

To understand the difficulties and the capabilities presently associated with solar observation, it is appropriate to review the instrumentation used for that purpose from the surface of the earth. An extensive review of solar instrumentation is included in this report as Appendix S-7. Some of the more critical points are collected in Section 4. **3. 3.**

4. **3.** 1 Engineering Considerations for a Solar Probe - In a study to analyze methods of obtaining information about the surface of the sun one necessarily must consider a close approach as a possibility of better observing its features and of making more accurate measurements of its physical parameters. However, one also realizes that there is some practical limit to this distance if the probe is not to be catastrophic. The various pertinent considerations were examined as part of our solar topography effort.

The first problem encountered, of course, is the orbital transfer requirements, especially the boosters that are necessary to achieve close approach orbits. Although out-of-the ecliptic orbits have excessive booster power requirements, close approach should be achievable in the ecliptic with the boosters predicted for the post-Apollo time period.

The necessity of power supplies for data transmission and equipment operation, requires a careful look at types available and the damaging influences they might be susceptible to, such as excessive heat or meteorites. Solar cells were examined in this respect.

A brief review of attitude control systems and the accuracies necessary for some missions was included because pointing accurately at some solar feature is desirable in many high resolution experiments.

Many of the adverse conditions of the solar environment were studied to specify the expected degrading influence of each on a solar probe vehicle. These included perturbations to its orbit and stability, radiation pressure effects, heating from intense thermal radiation (the major environmental effect), and communication difficulties arising from solar noise and occultation.

From all the considerations it was concluded that a perihelion distance of 0.05 astronomical units* for a solar probe vehicle may be attainable. However, it is cautioned that any particular mission would have to be evaluated in light of the instrumentation package it would carry before a practically achievable minimum perihelion distance could be defined.

For further details on this subject see Appendix S-5.

4.3.2 Cooperating Vehicles - From a study of the possible use of cooperating vehicles, it was concluded that they are not of value for the mapping of stable surfaces with the possible exception of the use of two side-looking radars where cooperating vehicles would permit measuring the distance between stations for additional control. However, such vehicles may find use in observing transitory phenomena. For example, because of

* An astronomical unit is the mean distance of the earth from the sun.

the rapid change of some solar phenomena (e. g. , granules), the slow rotation of the sun, and difficulty in close approach, cooperating vehicles would be required to observe these phenomena stereoscopically (see Appendix S-6).

Cooperating vehicles may also find application as relay stations for distant and difficult communications. For example, in Figure 3 there is illustrated two possibilities for the use of cooperating vehicles to aid distant communication. The upper illustration is a simple schematic to demonstrate a possible gain in received power with the same transmitter power distributed between cooperating stations. P_T represents transmitted power, G_t is transmitter antenna gain, A_r is receiving antenna aperture, R is a measure of distance, and P_R is the received power. In this simple case a factor of two is achieved, however, more analysis is required in terms of beamwidth and antenna size.

The lower diagram illustrates another possibility of using microwave feeders from distant probes to a well established laser relay station which combines the inputs and transmits on a single light channel. This would ease the beamwidth requirements for the transmitting probes and alleviate some of the bandwidth allocation problems of the home station. More analysis is also needed in this respect.

Cooperating vehicles could also insure the non-interference of antenna and sensor when both are required to point approximately in the same direction without a relay. In addition, the use of a relay station would permit acquisition of data from a solar probe when it was in eclipse, transit, or subtended so small an angle from the sun as seen from the earth that the earth receiving antenna would receive too much solar noise.

4. 3. 3 Solar Instrumentation - Before embarking on an extensive and expensive program of solar observation from space, it is a necessary preliminary step to review the capabilities and limitations of earth-based instruments to determine what, if anything, can be done better from a satellite, the moon or a solar probe. Particular attention must be paid to the size and weight penalties which generally go hand-in-hand with state-of-the-art terrestrial instrumentation.

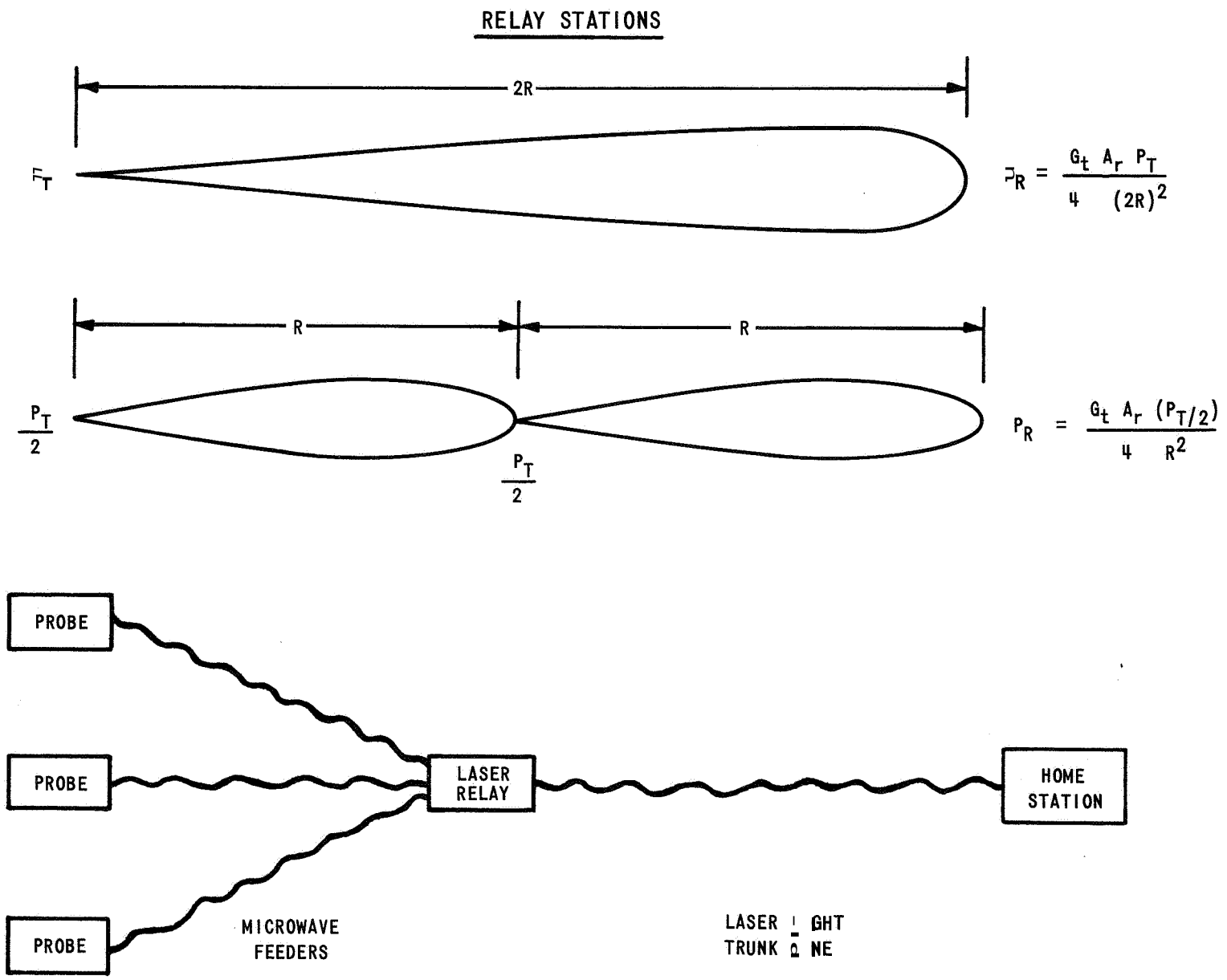


Figure 3 USE OF COOPERATING VEHICLES FOR COMMUNICATIONS

Telescopes have been treated at length to describe the critical considerations of these devices from mirror materials to apodization. Our review of telescopes included a brief account of the Stratoscope experiments (high altitude balloon mounted telescopes) which have produced the greatest resolution of the solar surface available, and a description of the results of a feasibility study for placing a one hundred and twenty inch telescope into orbit as an astronomical station above the atmosphere. The advantages and disadvantages of various types of spectrographs (prisms and gratings) used in solar spectral analysis have also been considered.

Instrumentation peculiar to solar observation that was also studied under this program included the spectroheliograph, the coronagraph and the magnetograph. The spectroheliograph is an instrument used to make photographs of the sun in monochromatic light of a single spectral line. The coronagraph is a device which allows the viewing of the solar corona without benefit of an eclipse by the use of an occulting disk to mask the much brighter photosphere. The magnetograph is an instrument which allows one to measure magnetic fields remotely by taking advantage of the Zeeman effect on spectral lines produced in the field.

The key result of such an investigation is the conclusion that more often than not the limit of earthbound observation of some critical feature on the sun is imposed by the atmosphere and not an instrumental deficiency. The practical resolution limits of ground-based solar observation are about 0.5 sec of arc for direct photographs and 0.8 sec of arc in the best spectra.[8] Improvements in these capabilities are expected to reveal new detail of the greatest significance. The size distribution of granulation, the fine structure of magnetic fields, the stranded structure of loop prominences, the local variations in line profiles, the turbulent velocity fields, and the minute structure in flares are noteworthy examples.

---

[8] REF. 8 - "Space Research-Directions for the Future", Report of a Study by the Space Science Board, Woods Hole, Massachusetts, 1965, Publication **1403,** National Academy of Sciences-National Research Council, Wash., D. C., 1966.

4.4 Experimental Aspects

Two experimental possibilities with application to future solar research have been examined during this program. The first potential experiment evaluated was the use of a laser probe to study the solar atmosphere by backscatter analysis. The results of this study are given in Section 4.4. 1. The detailed considerations are contained in Appendix S-8.

The other possibility is the use of Fourier transform spectrometry in space. The value of such interferometric devices in solar spectral analysis and the problems of implementing such instrumentation for use in space were assessed. The advantages and disadvantages are discussed in Appendix S-9. An abbreviated account of this investigation is presented in Section 4. 4. 2

4.4. 1 An Active Probe of the Sun - This section covers the results of a study and analysis of the feasibility of an active probe of the sun. It was hypothesized that a high power laser might be utilized to obtain information about the physical parameters measurable in the solar atmosphere by analysis of the reflected light. The mechanisms by which light could be reflected (backscattered) were examined to obtain quantitative comparison with the magnitude of the background radiation. Calculations using the best available values for the system constraints found such a system wanting by several orders of magnitude. It was concluded that a laser probe was not appropriate for such application. However, it was found that microwaves could probably produce useful information about the corona when used as an active probe in this manner. For more details see Appendix S-8.

4.4.2 Fourier Transform Spectrometry - Fourier transform spectrometry is a method of obtaining a spectrogram by harmonically analyzing the output of an interferometer. In some situations this method (as opposed to prism or grating spectrometry) can produce spectrograms with less noise, ease stabilization requirements, or provide data which may be more helpful in testing theories and models. However, calibration and ruggedness are two major problems faced by these instruments in space. At present considerable theoretical and experimental work is required to determine if the potential advantages are attainable in space and, if so, to secure them.

4.5 Recommendations

This section presents a compilation of the recommendations which are the result of our study of the measurement and description of the various physical aspects of the sun to better determine its dynamic topography. They are grouped according to several general areas of interest. The classifications appearing in parentheses after each title have been explained in Section 3.5 on page 29.

Solar Model:

1. Establish a NASA Model of the Sun by Reviewing Present Models and Defining Experiments and Analysis to Resolve Differences and Eliminate Unknowns (I,T) - Many investigators have constructed models of the solar atmosphere describing the profiles of temperature, pressure, electron density, etc. The disagreements in the results of these efforts are mainly caused by the major assumptions presently required for such modeling and the ambiguities involved in the data utilized as inputs to such models. The specifics of these two main categories require deeper analysis to define the requirements for space experimentation to eliminate the present unknowns. For example, the hydrogen-to-helium ratio, which is not accurately known at the present time, is just one of the inputs required for such a task. Perhaps a monitoring of the proton-to-alpha particle ratio of solar cosmic radiation produckd by flares could be utilized to refine the measurement of this quantity.

2. Incorporate the Influence of Rotation upon the Convection Zones in Future Studies of Solar Structure (N, T) - Present theories do not take into account the effect of the rotation that is known to be present.

Spectral Analysis:

3. Investigate Present Methods of Solar Spectral Analysis to Determine the Advantages to be gained by Escaping the Atmosphere, Getting Closer to the Sun, or Using Methods other than those Presently Used (I,T) - An important input to a solar model is the abundance of the elements. These are derived from the intensities and profiles of the spectral lines obtained from observations. These quantities are not known accurately enough at the present time.

4. Promote Basic Experimentation to Identify the Unknown Solar Spectral Lines (N, E) - Approximately thirty percent of the spectral lines found in the sun's radiation are as yet unidentified. It is believed that many of these lines are the result of radiation from highly ionized elements which have never been produced on earth. With the increasing availability of high energy machines, it may be possible to perform astrophysical experiments to duplicate the conditions of the unknown line generation, and thereby identify many of the unexplained lines. Feasibility study should precede experimental design.

5. Establish a Program to Compare the Intensities of Spectral Lines from the Sun above and below the Earth's Atmosphere to Determine more Accurate Transmission Characteristics (N, E & T) - Better knowledge of the transmission factor as a function of wavelength and time will be of value to earth-based observations in their subsequent data reduction.

Instrumentation:

6. Advance the State of the Art in Fourier Spectroscopy for Its Advantages in Solar Spectral Analysis (I,R) - A Fourier spectrometer has a distinct advantage over conventional spectrometers (e.g., a grating type) in sensitivity to weak lines and size. It has a disadvantage in spectral

resolution at the present time. Such a device could be improved for space application to great advantage. Several specific recommendations in this area are included in Appendix S-9.

7. Promote the Establishment of larger Optical Instrumentation in Space (N, R) (F,E) - State-of-the-art earth-based instruments are generally large. To make full use of the great potential of locating optical instrumentation in space requires that instruments of greater size than those lofted to date be constructed and maintained above the earth's atmosphere.

Probes:

8. Establish a Solar Patrol on the Side of the Sun not seen by the Earth to Monitor Sunspot and Flare Formation (F,E) - Asymmetries in the formational preferences of these phenomena create a great deal of interest in establishing a continuous watch on the entire surface of the sun.

9. Establish an Out-of-the-Ecliptic Satellite Orbit around the Sun to obtain Measurements from a more nearly Polar View (F,E) - It would be of great interest to astrophysicists to ascertain whether or not the sun's polar directed radiation is the same (in magnitude and color) as that measured in the ecliptic.

Granulation:

10. Obtain Greater Resolution of Solar Structure to determine if Micro-Structure Exists which is as yet Unknown to Us (N, E) - Present minimum sizes cited for cell size of photospheric granulation are at the resolution limit of the instruments recording them. The theories proposed to explain granulation could use such information.

11. Observe Solar Granular Field Shifts by Monitoring Motion during Growth and Decay to Obtain Velocities of Photospheric Rotation (F,E) (N, T) - The sun rotates differentially (faster at equator than poles). Measurements of the velocities by observing Doppler shifts at the limb and watching the movement of sunspots, differ. It is proposed to watch the

movement of granulation which is more characteristic of the photosphere than spots in hope of resolving this difference and obtaining more accurate values.

12. Perform an Observational Experiment (such as that proposed in Recommendation 23) to Establish or Refute the Presence of a Torsional Oscillation in the Sun Postulated by Richardson & Schwarzschild (F,E) - From an analysis of the rotational velocity of the sun over the solar cycle it has been theorized that a torsional oscillation exists superimposed on the differentional rotation. This would evidence itself as a deceleration and acceleration period coincident with the 11 year solar cycle and could be measured by accurate monitoring of the rotational velocities for some extended period.

Physical Parameters:

13. Move Close enough to the Sun and/or Obtain much Higher Resolution of the Sun's "Surface" to Allow Monitoring of a Single Valued Area as a Function of Time (F,E) - Present measurements such as a magnetic field temperature, etc. are all averages over areas having a distribution of values of the parameter desired. Such measurements would further the establishment of functional relationships necessary to establishing correct theories.

14. No Further Effort Should be Expended on the Use of A Laser as an Backscatter Probe of the Sun (I,T) - An analysis of the possibility of using a laser as an active probe of the sun to obtain parameter information by backscatter has been performed. The energy requirements far exceed the state of the art foreseeable in the near future.

15. Use a Microwave Probe to Profile Parameters of the Corona (I,T) (F,E) - Because the background magnitude is much less in the microwave than the optical region and because of the phenomenon of total reflection at the plasma frequency, it is suggested that an active probe of the sun using microwaves could produce valuable information about the structure of the corona.

Magnetic Field:

16. Make Magnetic Field Measurements of the Sun's Influence in Space External to the Earth's Magnetosphere (N, E) - Many of the theories concerning solar phenomena and energy transport require magnetic fields of certain magnitudes. Those which cannot be deduced from Zeeman splitting at a distance must be measured free of interfering fields. Coronal phenomena would become more amenable to accurate description with a greater knowledge of the sun's magnetic field extent and magnitude as well as variability.

5. APPENDICES

The appendices represent the reports of the detailed investigations which were conducted during the Second Phase of Project TECH TOP. They are broken into two main categories: those considerations which are primarily connected with a planetary topographical information collection system and those considerations which are applicable to the study of solar topography. They are presented in the order of their referral in the main text. The final appendix is a bibliography.

APPENDIX P-1

SYSTEM ANALYSIS FOR PLANETARY TOPOGRAPHICAL INFORMATION COLLECTION - ACQUISITION CONSIDERATIONS

By David A. Richards

This appendix contains a discussion of the detailed system considerations encountered primarily in the acquisition phase of planetary surface investigation. These include the problems attributed to planetary perculiarities and to the combination of elements (optics-sensor-readout) at the front-end of an imagery collection system. A qualitative run-down of the important considerations is followed by a quantitative analysis of some of the tradeoffs available within the constraints specified. This analysis is only a beginning to demonstrate what must be done on a larger scale (in amount of data points and complexity of functions) and should not be considered definitive.

I. INTRODUCTION

System analysis for optimal design proceeds through several stages from the original objectives as defined by the ultimate user and refined by the system designer, to the final design specifications. After the problem has been defined, one starts to hypothesize various system constructs which are candidates (or competitors) for its solution. Sometimes the concept to be employed is given with the problem. If not, each one listed undergoes a similar process. As we project the system concept into its necessary subsystems, we list the constraints (both physical and design) which limit some desirable facets for achievement of the objectives. Then, as we interrelate the subsystems, combined constraints are sometimes even more limiting and tradeoffs must be analyzed. The investigation continues as an iterative process by adding detail as the mathematical model becomes more complex. When system requirements, design, and state-of-the-art constraints have been thoroughly analyzed in combination, one can proceed to specify an optimal design.

An initial effort at implementation of such a procedure for a planetary topographical collection system analysis was begun on Project TECH TOP. The system construct was one which utilized stereo imagery. The acquisition portion of a total system was the primary concern. Analysis of the constraints imposed by the planetary environment and the tradeoffs available in the optics-sensor readout combination were as far as the investigation proceeded.

11. ACQUISITION CONSIDERATIONS

In the acquisition division of a system for the collection of planetary topographical information by means of stereo imagery, one can distinguish five subsystems for the purpose of analysis. They are the planetary environment, the observation platform, the sensor, storage and readout. The other divisions (transmission and processing) have been discussed in the main text (see Section 3. 2. 2).

Considerations that are important to the planetary environment subsystem are the illumination, the intervening atmosphere (if any), the orbit necessary and achievable, and the reflecting properties of the surface to be observed. The illumination can be natural or artificial. By natural we mean the irradiance of the sun which is a function of the orbit, the inclination of the rotational axis to the plane of the orbit, and the latitude, longitude, and local time of day of the area observed. When total area coverage is required, analysis must be made of the problems involved with dark area occurrence (e. g., arctic night on earth). Major portions of some planets are not naturally illuminated for extended periods. For example, the ninety-eight degree inclination of Uranus and its eighty-four year revolution period about the sun places major areas in darkness for approximately forty years (twice each revolution).* Artificial illumination requires an

* For the other planets, the data in Table I(P1) can be used to similarly examine illumination problems for dark area occurrence.

Table I(P1)

PLANETARY MOTION DATA

PLANET	SIDEREAL PERIOD	INCLINATION TO ECLIPTIC (DEGREES)	ROTATIONAL PERIOD	INCLINATION TO ORBITAL PLANE (DEGREES)
MERCURY	88 DAYS	7	59 DAYS	~7
VENUS	224.7 DAYS	3.39	~ 247 DAYS	— 6
EARTH	365.26 DAYS	0	23 HR. 56.07 MIN.	23.45
MARS	687 DAYS	1.85	24 HR. 37.4 MIN.	25.0
JUPITER	11.86 YEARS	1.03	9 HR. 55 MIN.	3.01
SATURN	29.46 YEARS	2.50	10 HR. 30 MIN.	26.7
URANUS	84.02 YEARS	0.77	10 HR. 50 MIN.	98.0
NEPTUNE	164.8 YEARS	1n79	16 HOURS	28.8
PLUTO	248.4 YEARS	17.2	6.4 DAYS	?

active source aboard the space vehicle and the power requirements may become excessive for orbital altitudes.

The atmospheres of many planets offer a major obstacle to surface observation. Their characteristics are summarized in Table II(P1). The Jovian planets are apparently covered by dense, cloudy atmospheres which would require radiation more penetrating than the visible spectrum. Most probably active microwave systems would be required. Extinction and turbulence effects require more definition from experimental results which might be obtained from precursor probes.

The orbit achievable about a planet depends upon the mass and radius of the primary, and the extent of the atmosphere, as well as our capability to inject the vehicle at the proper point and velocity. The necessary orbit is determined by a combination of things including area coverage, sensor parameters (focal length, sensitivity), the altitude, and the required resolution. "The arbitrary goal is to provide a system suitable for mapping at a scale of 1:250, 000. This scale is the minimum for topographic mapping, and the maximum in common aeronautical charts. It also is a reasonable value for the largest scale for initial mapping of the moon and terrestrial planets. "*

The reflecting properties of the planetary surfaces are not well known. The albedo and color index of the planet observed from the earth characterize the atmosphere more than the surface. The range of values to be anticipated for design of sensor dynamic range could also be derived from precursor probe experiments.

A listing of some of the planetary data mentioned above is given in Table III(P1).

The second subsystem to be considered is the platform **or** space vehicle from which our observations will be made. This platform must be stabilized to eliminate any excessive pitch, roll, **or** yaw that could degrade our results beyond our minimal requirements. The orientation and altitude

---

* (REF. 1) Levine, Daniel, "Radargrammetry", McGraw-Hill **Book** Company, Inc., New York, 1960.

Table II (Pi)
PLANETARY ATMOSPHERES*

PLANET	SUB-STANCE	DETECTED	AMOUNT cm atm (NTP)	BASIS OF ESTIMATE	REMARKS
MERCURY					PROBABLY FLUORESCING FREE RADICALS AND IONS PRODUCE HAZE
VENUS	CO_2	YES	10^5	SPECTROSCOPIC	MUCH BELOW THE CLOUD LAYER
	H_2O	YES	OCEANS	POLARIZATION OF CLOUDS	VAST OCEANS ARE ASSUMED TO EXIST BELOW THE CLOUDS
	N_2	YES	?	SPECTROSCOPIC	N_2 AND N_2^+ BANDS IN THE NIGHT SKY
	CO	YES	< 100	SPECTROSCOPIC	CO^+ BANDS IN NIGHT SKY AND ABSORPTION BANDS IN DAY SKY. LIMIT FIXED BY NEAR INFRA-RED BANDS
MARS	CO_2	YES	3600	SPECTROSCOPIC	
	H_2O	YES	?	POLARIZATION OF CLOUDS	POLAR CAPS CONSIST OF ICE. SOME CLOUDS HAVE THE POLARIZATION OF ICE CRYSTALS
	N_2	NO	1.8×10^5	TOTAL PRESSURE MEASUREMENT	N_2 IS ACCEPTED AS THE MOST LIKELY NON-CONDENSABLE CONSTITUENT
JUPITER	CH_4	YES	1.5×10^4	SPECTROSCOPIC	
	NH_3	YES	700	SPECTROSCOPIC	
	H_2	NO	2.7×10^7	DENSITY OF THE PLANET	ASSUMED TO BE PRESENT IN SOLAR PROPORTIONS RELATIVE TO METHANE ON THE BASIS OF THE MARCUS CALCULATIONS
	He	NO	5.6×10^6		
	N_2	NO	4×10^3		
	Ne	NO	1.7×10^4		
SATURN	CH_4	YES	35000	SPECTROSCOPIC	
	NH_3	YES	200	SPECTROSCOPIC	
	H_2	NO	6.3×10^7	DENSITY OF PLANET	ASSUMED TO BE PRESENT IN SOLAR PROPORTIONS RELATIVE TO METHANE
	He	NO	1.3×10^7		
	N_2	NO	9.5×10^3		
	Ne	NO	2.7×10^4		
URANUS	CH_4	YES	2.2×10^5	SPECTROSCOPIC	
	H_2	YES	9×10^6	CALCULATED ON HERZBERG'S ASSUMPTION	He AND H_2 ARE ASSUMED TO BE EFFECTIVE MOLECULES IN PRODUCING TRANSITIONS OF H_2
	He	NO	2.7×10^7		
	H_2	YES	4.2×10^6	CALCULATED ON UREY'S ASSUMPTIONS	N_2 AND H_2 ARE ASSUMED TO BE EFFECTIVE MOLECULES IN PRODUCING TRANSITIONS OF H_2. SOLAR PROPORTIONS OF He AND H_2 ARE ASSUMED
	He	NO	8.6×10^5		
	N_2	NO	4.2×10^6		
NEPTUNE	CH_4	YES	3.7×10^5	SPECTROSCOPIC	
	H_2	YES	LARGER THAN IN URANUS	SPECTROSCOPIC	ASSUMED TO BE EFFECTIVE MOLECULES IN PRODUCING TRANSITIONS OF H_2
	N_2	NO			SOLAR PROPORTIONS OF H_2 AND He ASSUMED
	He	NO			

* REF. 2 - UREY, HAROLD C., "THE ATMOSPHERES OF THE PLANETS," HANDEUCH der PHYSIK, VOL. LII, ed. S. FLUGGE, SPRINGER-VERLAG, BERLIN, 1959. p. 413.

Table III (PI)

PLANETARY DATA

NAME	DISTANCE FROM SUN (A.U.)	MASS (M_e)	RADIUS (KM)	LIGHT (L_p/L_e)	ALBEDO	ATMOSPHERE
MERCURY	0.387	0.0543	2480	6.7:1	0.055	NEGLIGIBLE
VENUS	0.723	0.8136	6200	1.9:1	0.80	DENSE
EARTH	1.00	1.0000	6371	1:1	0.35	TRANSPARENT
MARS	1.524	0.1080	3310	1:2.3	0.15	TRANSPARENT
JUPITER	5.203	318.35	70,000	1:27	0.44	DENSE
SATURN	9.539	95.22	57,500	1:91	0.42	DENSE
URANUS	19.19	14.58	25,500	1:370	0.45	DENSE
NEPTUNE	30.07	17.26	24,800	1:900	0.52	DENSE
PLUTO	39.46	0.93	2900	1:1560	0.16	UNKNOWN
MOON	1.00	0.0123	1738	1:1	0.07	NEGLIGIBLE

1 ASTRONOMICAL UNIT (A.U.) = 1.496×10^8 km.

1 EARTH MASS (M_e) = 5.976×10^{24} kg.

L_p/L_e IS THE RATIO OF ILLUMINATION AT A PLANET TO THAT AT THE EARTH-

of the station must also be determined for each observation to permit data reduction. These operations admit of several different solutions and the candidates must be evaluated in terms of cost-effectiveness with respect to the entire system. For example, a star camera used to determine attitude must be analyzed to include the attendant data reduction problems and accuracies. Altitude can be measured by optical or microwave radar. The advantages and disadvantages of these elements of the platform subsystem must be compared for the particular application at hand. Microwaves would be preferred for dense atmosphere penetration. However, lasers are more accurate when the atmosphere is not a serious limitation. The platform is also limited in size and weight achievable and/or desirable.

The sensor is an extremely critical subsystem because it defines many of the requirements imposed upon other subsystems. If the sensor choice is photographic film, a means of development must be supplied and the readout subsystem receives the constraint of a silver grain image as its input rather than a charge pattern characteristic of a television sensor. The sensor also imposes its sensitivity and dynamic range upon the system. Sensor considerations include limitations on the signal power collector (antenna or aperture), filtering components, the compatibility of the sensitive element with the launching and space environment, and the integration time of its response. The specification of the sensor subsystem generally defined the nature of the entire system. (Sensors have already been considered during the First Phase. *)

The fourth subsystem that has been delineated in the acquisition division is the storage phase. When necessary, usually because of transmission constraints, it is accomplished by several different methods dependent upon the sensor output. When the sensor produces an electrical signal as its output, magnetic recording (tape or core) is appropriate. Images

---

* (REF. 3) Campbell, Charles E., "Project TECH-TOP - Study of Lunar Planetary and Solar Topography", First Phase Report, Cornell Aeronautical Laboratory, Inc., Report No. VC-2104-D-1, August 1966.

can be stored as silver deposits in a developed photographic film or as charge patterns in electronic storage tubes. Some data processing could also be included in the storage function if it could be done to some advantage with respect to its cost.

The readout process is the final subsystem in the acquisition division. For imagery some form of scanning is generally required. This can be mechanical, optical, electronic or a combination of these. There are various patterns and speeds of scan that have been devised and used. The most popular scanning method is the parallel raster scan used, for example, in commercial television. The alternative to scanning is the use of many channels simultaneously (one for each element producing information). The signal that is ultimately fed out of the acquisition division to the transmission division may be in analog (continuous) or digital (discrete) form.

A thorough analysis of all these considerations requires resort to mathematical description. Limits must be obtained as numerical values from the state-of-the-art in each particular instance and interrelationships must be characterized by functional connection of the critical parameters. When such numbers and functions are available, tradeoff analysis can be efficiently accomplished by machine computation. When such a procedure is followed, weakest links in the chain from source to user are more readily pinpointed for research and development requirements to promote advancement of the system state of the art.

III. OPTICS - SENSOR - READOUT ANALYSIS

A parametric analysis of some of the important relationships of the optics-sensor-readout combinations was performed in an elementary numerical form. A more detailed analysis would proceed to compile transfer functions of the various subsystem elements (as described in Section 3.2.3 of the main text).

The first relation to be considered is the angular resolution requirement as derived from the ground resolution desired by the user, and the orbital altitude achievable. This is simple given by:

$$S = h\theta \qquad \text{(P1-1)}$$

where S is the ground resolution defined as the minimum distance between *two* adjacent features or the minimum size of a feature on the ground, h is the altitude, and θ is the required angular resolution. (We have assumed high contrast objects. **A** more intricate analysis is required to include the range of contrasts that may be encountered.)

Once the angular resolution requirement has been established, a choice of collecting aperture for the sensor can be specified if the wavelength of the radiation to be utilized is chosen. These quantities are related by the Rayleigh criterion

$$\theta = \frac{1.22\lambda}{D} \qquad \text{(P1-2)}$$

where θ is the angular resolution obtainable with an aperture of diameter D illuminated by radiation characterized by wavelength λ . (We use a mean wavelength here. The actual case requires consideration of the multiplicity of wavelengths present. Also, super-resolution beyond this Rayleigh limit is possible in special cases and should be studied further for space applications,) The results of calculating θ for wavelengths from ultraviolet to infrared for a range of aperture sizes are presented in Figure 1(P1).

After obtaining a radiation collector consistent with the initial requirements, we must now consider the sensitive element of the sensor package. The resolution capability of this device combined with the angular resolution requirement define the focal length necessary for the collecting optics according to the equation

$$f = \frac{l}{r\,e} \qquad \text{(P1-3)}$$

where f is the focal length, r is the sensitive element resolution capability and θ is the angular resolution requirement obtained above. (The angular resolution capability of an optical system off-axis is generally

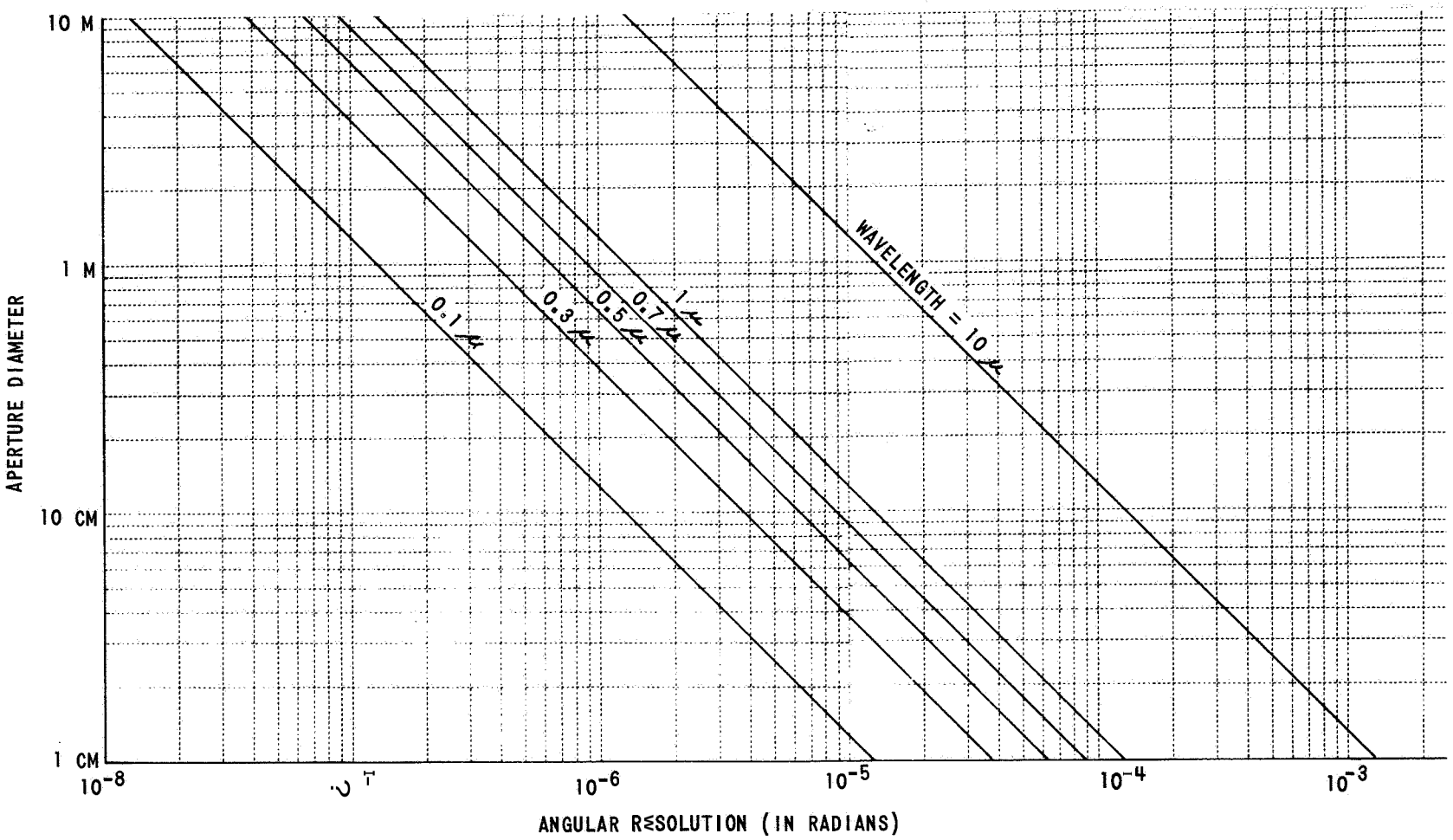


Figure 1(P1) APERTURE LIMIT

not as good as that on-axis. This results in the resolution at the center of the format being greater than that at the periphery. This has to be taken into account in a more complex manner.) Figure 2(P1) gives the angular resolution obtainable for various combinations of focal length and sensor resolution capability.

The focal length and aperture diameter define the f-number or focal ratio which is important in energy considerations and consequently the sensitivity requirement. These considerations are not examined here.

To convert a two-dimensional image into a one-dimensional time varying signal for transmission generally requires some form of scanning readout (flying spot scanner for photographic images or electronic scanning for television images). The form of the scan most often utilized is a parallel raster as used in commercial television. The resolution which a readout parallel scan can preserve is given by

$$r = \frac{L}{2F} \qquad \text{(P1-4)}$$

where L is the number of lines in the raster and F is the magnitude of the format dimension in the direction perpendicular to the raster. (The actual resolution a scan system can preserve involves a more complex analysis and the possibilities of other scan system applications must be explored).

Another factor which can limit the resolution capability of the acquisition subsystems is image motion blur. The results of some computations representative of the image motion that would have to be compensated for the oribital motion experienced during a range of exposure times for the planets and an altitude of 100 kilometers, are presented in Figure 3(P1).

Whenever a calculation indicates a requirement which exceeds the state of the art, one must return to the previous calculations and adjust any parameters that can be made compatible with a reasonable continuation, or relax the initial requirements. Of course, any calculated parameter of a subsystem component is always at the mercy of considerations such as those presented in Section 11. For example, although some high resolution devices

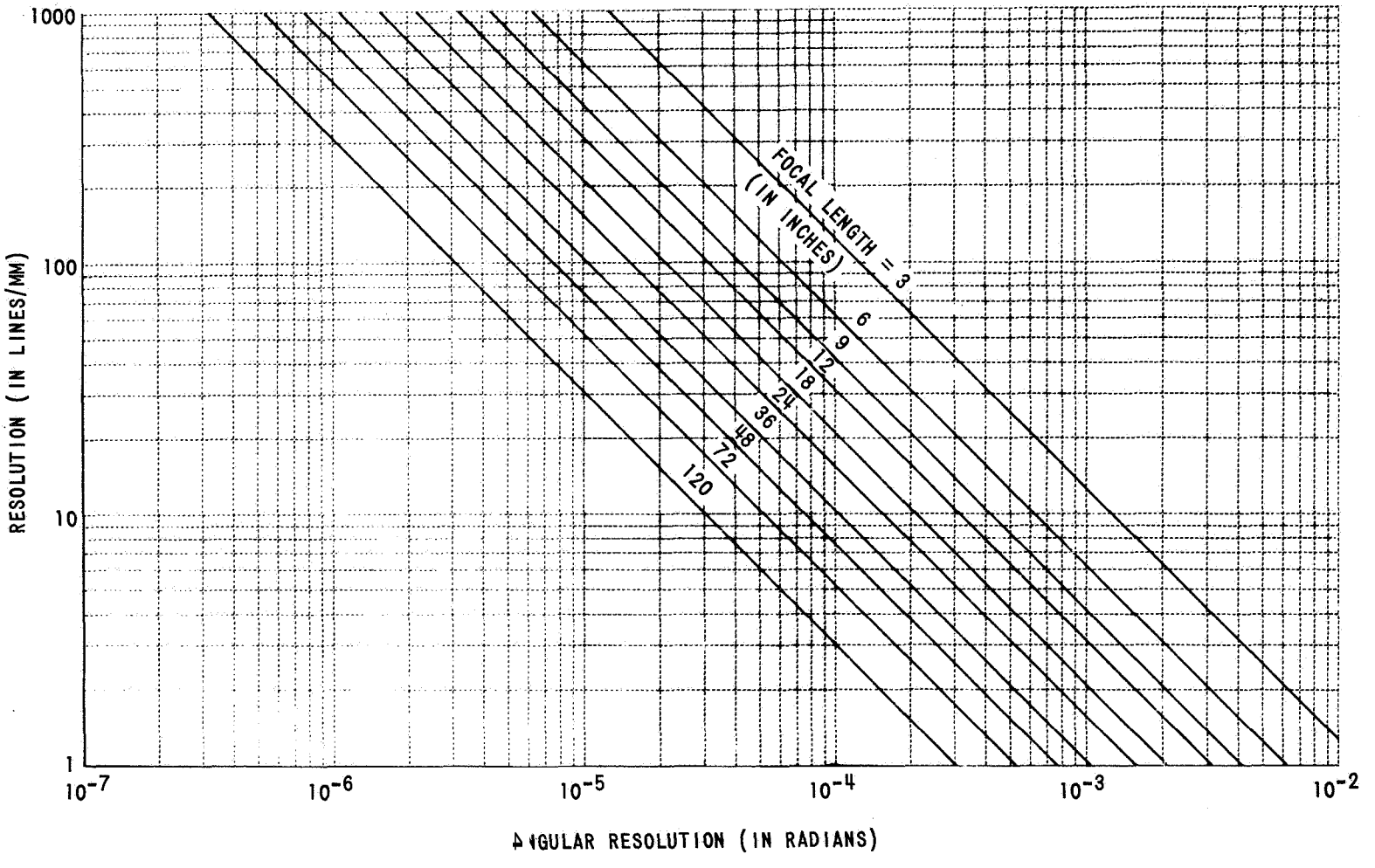


Figure 2(P1) CAPABILITY OF DETECTOR

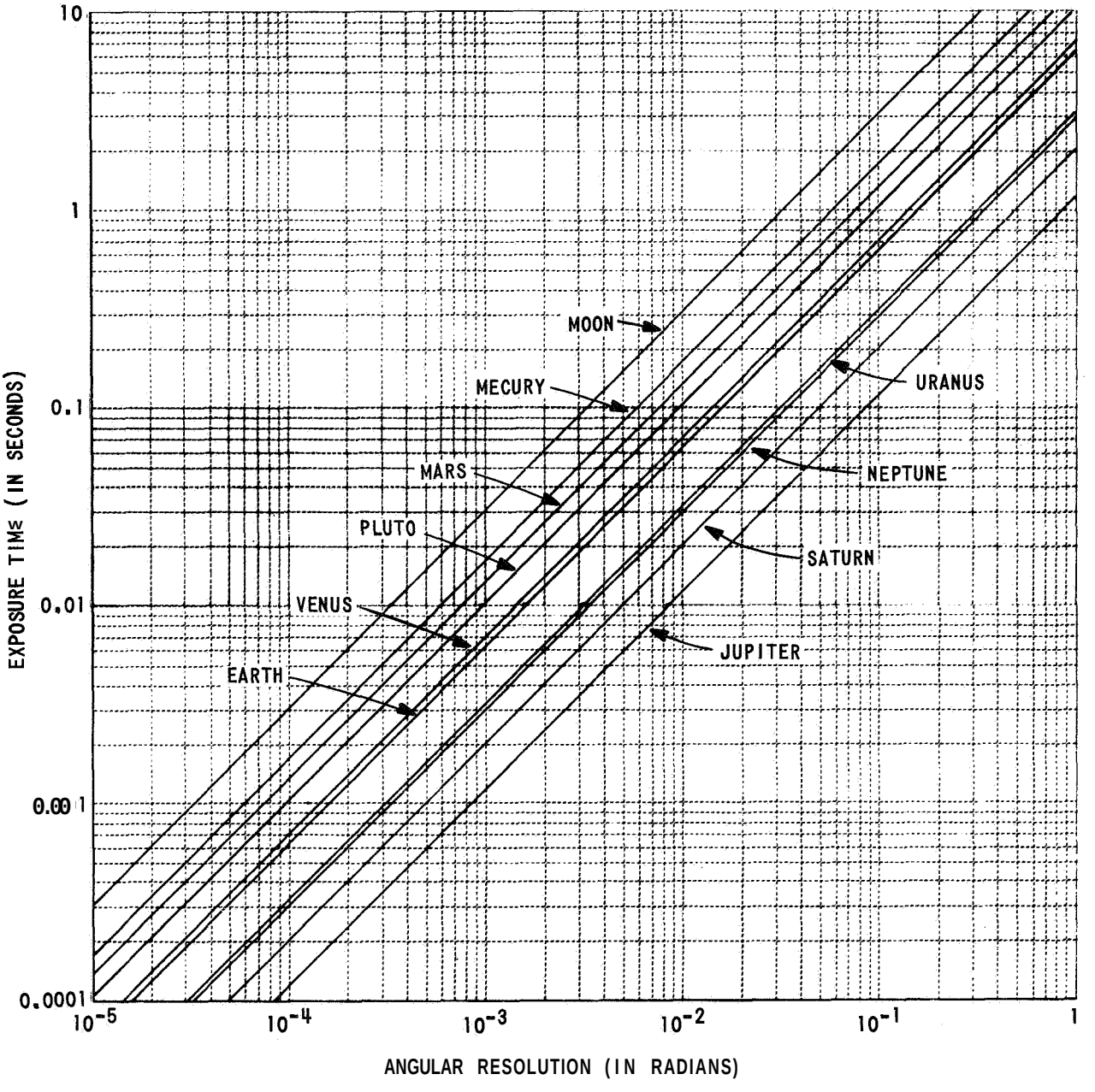


Figure 3(P1) ORBITAL MOTION LIMIT (100 KM ALTITUDE. NO IMAGE MOTION COMPENSATION)

represent the state of the art in this respect, they may not be compatible with the launch and space environment, and these constraints may force the choice of a sensor of less resolution than desired and available.

IV. CONCLUSIONS

The complex task of obtaining topographical information about the planetary surfaces from space vehicles by means of stereo imagery involves the interplay of many quantities and components. For the optimal use of time and resources and for the achievement of optimal results, trial and error cannot be tolerated. Guesswork must be minimized. It is concluded from the foregoing considerations and analysis that a more detailed and thorough system analysis is required for the specification of an optimally designed project to accomplish space objectives. A quick-and-dirty cementing together of off-the-shelf items simply to meet an arbitrarily established timetable is ultimately wasteful of both time and resources. The carefully studied and intelligently executed program will obtain more results and they will be of greater value to the ultimate user.

All these calculations are merely preliminary to a more rigorous generation of an optimal design by mathematical system analysis. They are rendered here to give some indication of the procedure and difficulties.

APPENDIX P-2

IMAGE PROCESSING WITH RESPECT TO SPACE TOPOGRAPHY

By John D. Gallatin

This appendix discusses various types of image processing from the standpoint of space topography, pointing out which are useful, which are not useful, and those for which further research is required to determine if they can be made useful.

I. INTRODUCTION

Image restoration is the process of obtaining from an actual image, obtained under various degrading influences, an image which is improved in some respect. Thus, there are several types of image restoration corresponding to several ideals of image improvement. Because it may not be possible or necessary to obtain a perfect image, but useful to recover one in some sense better than the degraded image, image processing rather than image restoration is considered.

Stereo photogrammetric topography requires the measurement of stereo parallax and the interpretation of detail. In this appendix, image processing is considered in terms of these objectives.

Properties of images considered for processing are the gray scale rendition, the location of sub-images, and the rendition of detail. The detail processing is further divided into stripping the effects of sampling, suppression of unnecessary information, general purpose homogeneous image enhancement, and inhomogeneous image enhancement. By homogeneous processing, we mean that the operations performed on the degraded image are independent of position, and by inhomogeneous, that the operations performed do depend on position. There is also the concept of quasi-homogeneous processing, where the operations change so slowly with position that they may be considered constant over the size of detail they are designed

to process. Although sampling, stripping, and suppression are special cases of homogeneous processing, they are sufficiently specialized in both method and motivation to require separate discussion.

The remainder of this appendix contains sections discussing the motivation for various types of image processing, the methods used in processing, and the methods used to determine the operations to be applied in processing.

11. MOTIVATION FOR IMAGE PROCESSING

Not all types of image processing are helpful in attaining the final objective of a map. Thus, in this section, various types of image processing are discussed from the standpoint of applicability to space topography.

The recovery of gray scale, or tone restoration, is not usually expected to be of value for interpretation and compilation. Both human and automatic compilers operate satisfactorily with the distorted gray scales presently available. Interpreters can even use the reversed gray scale of photographic negatives. The above statements assume that for human inspection, the detail being inspected is at the proper brightness level with respect to its surrounding to be optimally visible. If this is not true, then this may be accomplished by masking or by gray scale restoration, or perhaps even predetermined distortion of the gray scale. When accurate recovery of photometric information is required, the corrections may be made after the brightness or density or whatever of the sub-image has been measured. This may usually be done at no added inconvenience as there are many other computations to be made in photometric situations.

The correction of distortion has obvious application to mapping. The only question is whether the map is best achieved by correcting distortion in the image, at the time of, or after compilation.

Removal of spurious structure, such as half-tone and scan structure, is sometimes required in order to prevent troublesome moiré effects in stereo interpretation and human compilation and to prevent spurious results in certain types of automatic compilation.

The use of other types of homogeneous restoration is open to question. It has been shown (see Appendix **P-4)** that purely homogeneous restoration is of no value in automatic stereocompilation. It is possible that this result also applies to manual compilation. It is not certain that suitable quasi-homogeneous processing is practically realizeable. It is possibly that homogeneous and/or quasi-homogeneous processing may be of value in facilitating the interpretation of imagery, but because this would depend on the skill of the interpreter, the sub-image being sought and their backgrounds, and the degradations to which the imagery had been subjected, no definite answer to this question has been given.

The capabilities and limitations of inhomogeneous processing have not been sufficiently investigated to permit statements regarding its applicability to space topography to be made. Only recently have investigators considered noise with respect to inhomogeneous processing. It was found that some processes are extremely sensitive to noise. As research on inhomogeneous processing continues, it is to be hoped that more refined inhomogeneous methods will be discovered which will keep the noise down to tolerable levels at only a reasonable cost in the recovery of the image. For example, the author has found that by using a suitably chosen sampling interval, the noise involved in the removal of image motion can in some cases be limited. Thus, more must be known about inhomogeneous processing before its capabilities in space topography can be stated.

Suppression is any process in which unnecessary information is removed in order to ease requirements on subsequent data stores and/or processors. A good example of this is the automatic exposure control camera. If the camera does not record the lens stop and exposure time, one cannot determine the absolute radiances of objects in the scene, but this information is often not required.

III. METHODS OF AND DEVICES FOR IMAGE PROCESSING

A) Methods for Restoration of Gray Scale

The—storation of gray scale can be done by the use of photographic masking techniques, by the use of gravure film, by suitable electronic analogue scanners or by digital computation.

Both photographic masking and the use of gravure film are explained for use to increase ccntrast in the parts of the image with least exposure; that is, on the toe of the H & D curve. They may be used to increase contrast on the shoulder by operating on a positive copy instead of on a negative. Masking is done by making a high contrast print of the negative using only enough exposure in the printer to cover the toe of the curve. Then, this mask is either reversal processed or copied to get back to the same contrast sense as the negative to be improved. When the mask is placed in contact with the negative, the composite image will have higher contrast for the regions of least exposure, thus matching the contrast of the regions of greater exposure. Gravure film is made with the proper grain size distribution to exhibit higher contrast on the shoulder of its H & D curve, which correspond to the toe region for the negative.

Analog electronic circuits to restore gray scale have been used for many years in broadcast television, Because the brightness of a kinescope increases more rapidly than linearly with respect to grid voltage, and because not all television pickup tubes are linear, circuits called rooters are used to pre-distort waveforms between the pickup and the kinescope. The principle of operation of one type is to increase the load on its source as the signal level is increased. This is accomplished by the use of series diode-resistor combinations which are backbiased from a power supply for small signals and which shunt away part of the large signals.

Finally, the gray scale may be restored by the use of digital computation. The image must be available as numerical samples. Then, the original density or radiance may be found by tabular look up and interpolation techniques.

B) Methods for Restoration of Distortion

One must distinguish between two types of distortion. The older and more familiar type results in the displacement of sub-images from their idealized locations as a function and is independent of object radiance. Examples of causes of this type of distortion are the lens aberration called distortion, and defective fabrication of the lens and camera. When destructive readout is used, such as in television camera tubes, a new type of distortion arises. In vidicons and orthicons, more of a electron beam is required to neutralize charge in the brighter portions of the image. Since the charge neutralizing process starts at the leading edge of the scanning beam, and continues until completion, it is seen that the center of the effective part of the beam is near the leading edge of the beam for images of dim objects and shifts back relative to the beam for brighter ones. Thus, this new distortion depends on radiance and is independent of location when the entire format is scanned by an ordinary raster scan.

Not much is known about this destructive scan distortion. Therefore, to the author's knowledge, no methods of correction have been proposed. Fortunately, when both images of a stereo pair are subjected to raster scan in the same direction, to the first order, the errors created are horizontal and not vertical; that is, spot altitudes will be correctly measured, but plotted on the map at the wrong spot. Because horizontal errors are often less serious than vertical errors, this destructive scan distortion may not be a serious problem.

Methods for the correction of the older type of distortion are well known (see Ref. 3, P. 111ff). They include the Porro-Koppe reverse distortion method, methods involving the use of specially figured plates,

methods involving the use of mechanical motions in printers, interpolation methods for images represented by digital samples, corrections applied to the compilation mechanism, and corrections applied after compilation.

The Porro-Koppe principle is that the distortion could be completely compensated by using the camera lens in the plotter. This includes only distortion caused by the lens and not that caused by improper orientation of the plate or film. Also, the reverse distortion principle does not include the fact that camera operates at infinite focus, but the plotter operates over a range of close image distances. This requires either that the camera lens be used at a different set of conjugate distances, or that it be supplemented with a close-up lens, and that it be stopped down. Both of these operations can change the distortion. The reverse distortion method is not particularly applicable to space topography because the camera lens may not be recovered in a serviceable condition or at all. Also, as lens distortion is at least partially caused by defects of manufacture, and may be changed by the shock of launch, it is not considered practical to attempt to duplicate it. Hence, other methods are required for correction of distortion.

Distortion is sometimes compensated by the use of specially figured glass plates. These are inserted into printers (conceivably they could also be used in plotters) and refract the image-forming light according to local thickness and wedge angle. These are most useful for correcting purely radial distortion as it is very difficult to fabricate them for general distortion. Also, depending on the size of the negative and the sizes of the distortions, and whether the printer lens was designed for use with them, they may reduce image detail quality.

It is also possible to correct distortion by use of mechanical motions in a projection printer, The principle is to print only part of the negative, then to move the lens and print another part. The process may be carried out either in steps or continuously. If the process is done in

batches, the correction will be complete only at a few points or along a few curves for each batch. If the process is continuous, then there will be image motion during the exposure. Also, generation of cams to supply the motion can be just as severe a problem as for mechanical correction during plotting, which is described below.

If the image exists in the form of digital samples, then correction for distortion may be accomplished by use of a digital computer. One has available a set of image values (radiances, densities, etc.) at presumably locally equally spaced sample points, but locally these sampies were taken at the wrong points. One requires a new set of image values at the right points. This is seen to be an interpolation problem. The diffraction theory of optical imaging guarantees that the image values will be band-limited with respect to spatial frequency. Hence, the interpolation may be performed by the use of diffraction functions; i. e., $\frac{\text{Sin } x}{x}$ as outlined in discussions of the Shannon sampling theorem (Ref. 1, 2), if the samples are adequately closely spaced. The interpolation is theoretically without error or loss in detail but requires image values extending off to infinity in both directions. Practically, there may be a small loss from termination of the interpolation formula.

Various methods may be imagined to use distortion data to correct maps at the time of compilation. An example of mechanical correction is provided by the Kelsh plotter. Here, a ball-shaped cam moves the lenses in the projectors to correct the radial distortion. The cam is moved to correspond to the part of the model being compiled. The correction applied is applicable at only one image distance. For satellite photography, this is a reasonable procedure because the relief is small compared to the altitude. In other cases, mechanical linkages may be devised to make the correction hold over a range of image distances. Although in the actual Kelsh plotter only radial distortion is corrected, it is obvious that all types of distortion may be corrected by mechanical linkages. If digital techniques are in use, it may be possible to make the correction by computation.

It is also possible to correct the distortion after the map is compiled. To do this, one must know the horizontal and vertical errors caused by the distortion. Then, the horizontal errors may be corrected by shifting all marks on the map by the required amount, and spot altitudes may be corrected simply. The contour lines may be corrected by shifting a distance corresponding to the error divided by the terrain slope. This method is not recommended because additional work is required to find the errors caused by the distortion, and because it can be inaccurate on contours as the slope is no longer accurately available.

C) Methods for Detail Processing

Methods available for general homogeneous processing are all equivalent to cross-correlation, but there is a special device which can only be called an optical cross-correlator. Other methods of general image restoration rely on coherent optical processing which is sometimes called coherent optical spatial filtering, and sometimes called holographic image restoration. Processing by digital computer is general in the sense that it may be used to cross-correlate with an arbitrary function, but the image must be supplied as samples and further processing is required to strip off the sampling and restore a continuous image. There have been several image processing machines which operated by means of an analog scan of the image; they are limited as to the function the image is correlated with. Finally, there are devices which can correlate with only one function, these are dodgers, used only for suppression, and the "incoherent"* halftone strippers used only for removal of sampling structure.

1 Detail Processing by Incoherent Cross-Correlation

Figure 1(P2) shows one of the possible configurations of cross-correlation. Light from the extended diffuse sources illuminates

---

* These devices are used in an incoherent system, but an understanding of their operation requires physical optics.

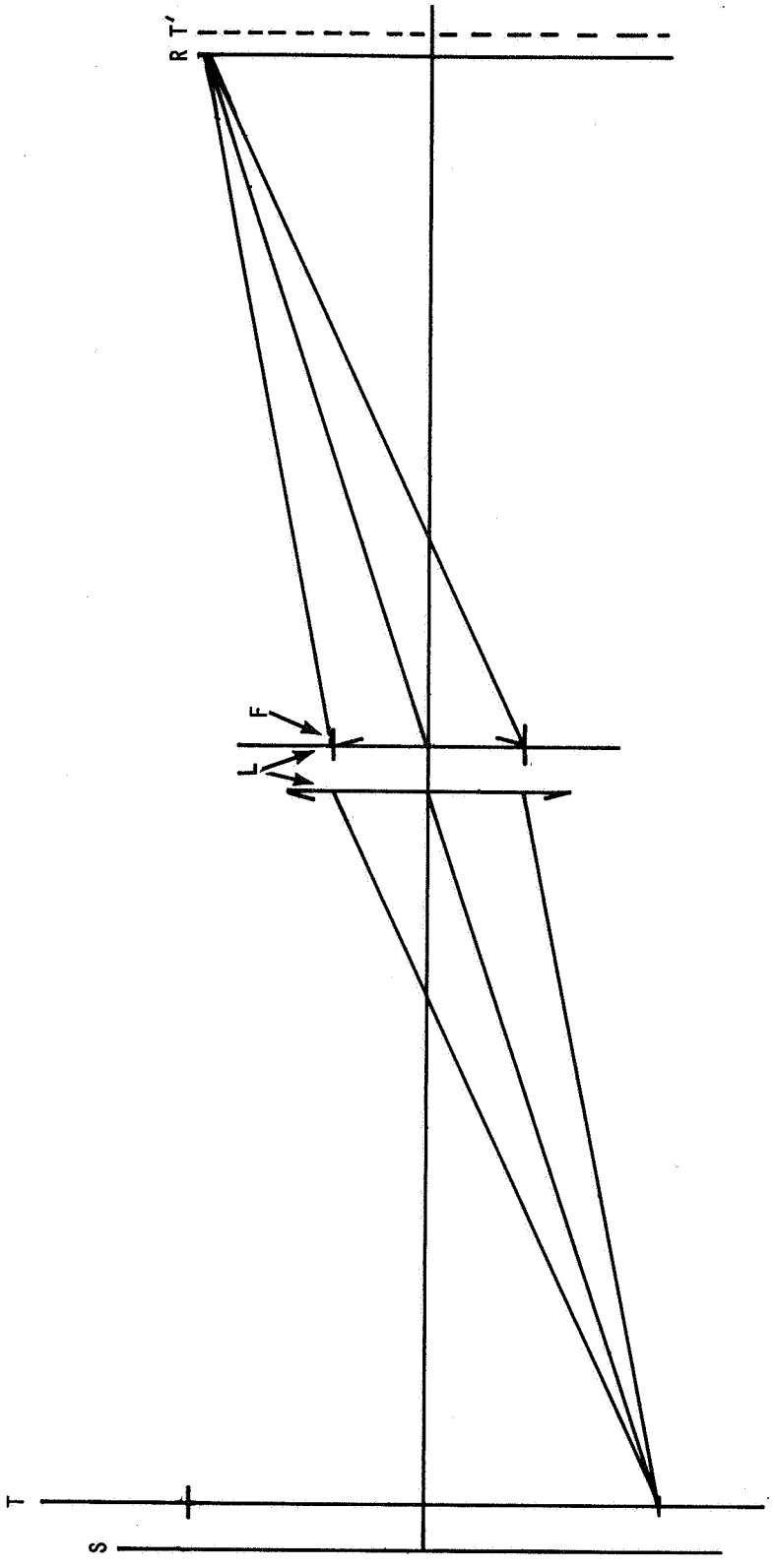


Figure 1 (P2) OPTICAL CROSS-CORRELATOR

the transparency to be processed, T. This transparency is imaged by the thick lens L (represented by its principal planes) on the conjugate surface, T'; however, the defocused image is intercepted by the film or other receiver at R. The system is shown to operate at essentially unit magnification, this is not necessary; also one may use the defocused image behind the surface T'. The filter F is the exit pupil of the lens. It is not essential that the filter be at the second principal plane, but it is easier to see the scale of the various functions involved if it is. The processing is based on the assumptions that light travels in straight lines and that the total amount of light at any point in the plane R is the sum of that carried by all of the rays from the filter to that point. Each ray's energy is proportional to the product of the filter transmittance at the point of the transparency toward whose image it proceeds. Then, when scale is taken into account, the irradiance in the plane R is seen to be a correlation of the transparency and the filter.

Other types of correlations may be constructed in which the filter is placed near the source, and the receiver is placed slightly behind the transparency. No lens is used. This correlator operates by the same principles as those above. Reference **4** contains a description of several correlators and the fluorescent screen process.

The above correlators operate by using a transmittance filter and defocus. Another type, seldom used, has a filter consisting of a specially figured plate as the exit pupil of the lens designed to spread the light from a point into the desired function. Determination of the optimum plate figure from the desired correlation function is not completely unaerstood. A non-optimum filter may be designed by using the facts that a cylinder generates a straight line, a cone generates a ring, and a sphere (defocus) generates a solid circle. The figure would be determined by piecing together these simple elements. It would, however, be very difficult to fabricate.

A practical limitation of incoherent cross correlators is that the correlation function usually has negative parts, and that no method has been proved suitable for simulating negative light. Methods tried include the use of the Herschel reversal in film, quenchable phosphors, photographic division (19) and electronic methods. For the use of the Herschel effect and phosphor quenching, the positive regions of the filter passed only blue light and light of shorter wavelengths, exposing the film or exciting the phosphor, and the negative regions passed only red and infrared, reversing the film or quenching the phosphor. The output was obtained by developing the film or by viewing or photographing the phosphor screen through a green filter, which passed only the light the phosphor radiated. It was found that neither of these methods was linear over an adequate range. Photographic division was performed by making two correlations, one with the positive parts of the aperture filter, and one with the negative parts. In each case, the other parts of the aperture filter were opaque. A photographic reversal was made of one of the correlations, and the output was obtained by viewing the reversal and the other correlation in contact. This process is called photographic division because the operations correspond more closely to dividing the positive part of the correlation by the negative part than the subtraction implied by the definition of correlation. Better results were obtained by this method than by the previous two. The most serious limitation of this process was again obtaining linearity over adequate range. It is difficult to make a copy at the exact contrast required, and even then the contrast will not be proper over the entire range. There was also a slight registration problem caused by film shrinkage and printer geometry; presumably, this could have been remedied by the use of glass plates and a different printer layout.

Recently it has been proposed to use an electrostatic imaging surface such as in "Xerox" copiers, as the image receiver, to perform the two correlations sequentially, and to obtain the effect of

subtraction by change in voltage applied to appropriate leads. This proposed process is evidently free of registration problems. Hopefully, by the use of sufficiently high voltages, linearity could be obtained over an adequate range.

Many methods can be imagined to convert the correlations into electrical signals so that the difference may be taken electronically. Among the best of these methods is one using polarization coding and sampling which is discussed below [20]. It may be implemented using either a camera tube as the receiver or by reversing the system replacing the camera tube by a flying spot scanner and the source by a phototube. The camera tube method is perhaps more easily understood in view of the previous discussion and is the method described below. However, the flying spot scanner method permits better correction of distortion; first, because the scanning is non-destructive, and second, because it is possible by means of a beam-splitter to use the same flying spot both to perform correlations and to print the result so that the correlation process will not introduce unknown distortion from changes in registrations. Hence, the flying spot scanner method is more applicable if the imagery is to be used for parallax measurement.

In the camera tube mechanization, the positive parts of the aperture filter pass only one {polarizationform and the negative parts pass only a form orthogonal to it. An example of orthogonal polarization forms are two linear polarizations at right angles. A coarse grating, with stripes perpendicular to the direction of scanning, is either imaged on or contacted to the tube face. Every other stripe of the grating passes only one of the polarization forms of the aperture filter and the remaining stripes pass only the orthogonal form. Thus, the irradiance pattern on the tube face consists of alternating strips of the positive and negative parts of the cross correlation.

When the pattern is scanned, the current (assuming gray scale restoration, if necessary) consists of pulses, every other one proportional to the positive part of the correlation, and the remainder proportional to the negative part. Assuming that synchronism could be maintained, the correlation could be recovered by synchronous detection followed by low pass filtering. However, if the correlation is always positive, as it should be for image detail processing, then a simpler mechanism is possible. The video current is divided into two parts. One part is delayed by the time required to scan one stripe width, and then subtracted from the other. The correlation may be recovered by half-wave rectification.

Closer inspection of the above process reveals that the resulting electrical signal is not exactly proportional to the cross correlation, but this can be remedied easily. First, by a proper shift of the negative part of the aperture filter with respect to the positive part, the difference may be taken between correlation values at the same point. Second, it is seen that the process generates samples of the correlation function. However, if these samples are taken at a close enough spacing and then properly smoothed, no loss of information need result.

This process solves the negative light problem, but at the sacrifice of some of the advantageous characteristics of cross correlators. First, because it is a sampled system, it may be necessary to remove the effects of sampling in some cases. The seriousness of the sampling is less for this system than for digital processed, where it is discussed, because the scan direction may be changed to avoid noise effects more easily than for digital processes and because any low-pass filtering required at the input may be incorporated in the aperture filters. Second, because it is a television system, and TV systems cannot cover as large a format at the same resolution as photographic systems, in some cases it would be necessary to process in batches. This is aggravated by the requirement that the resolution of the TV system be about twice that of the optical system to provide for sampling.

A second problem of cross-correlators is the determination of the aperture filter. In the previous discussion, geometrical optics was used to derive that the reference functions could be taken approximately as the shadow image of the aperture filter on the receiver as cast by a point source in the object plane. Hence, one may fabricate an approximate aperture filter by proper scaling of the reference function. A rigorous formulation of aperture filtering would be based on the diffraction theory of aberrations. Solution for the aperture filter(s) within this formulation is very difficult, and it is not certain that a reasonable filter(s) can be obtained within a reasonable amount of time, compatible with use of a wide range of wavelengths. The assumption of a source uniform over area and angle can be satisfied by use of proper equalization masks and lenses to generate the uniform source.

2. Detail Processing by Coherent Apparatus

Figure 2 (P2) shows the optical layout of a holographic or spatial filtering image restoration system. Systems with fewer lenses are possible, but the lenses would be larger. Use of more and smaller lenses reduces difficulties in obtaining lenses free of defects, mechanical problems, and cost, and provides more opportunities for the correction of aberrations. The liquid gates are not universally believed to be required in all cases.

The purpose of liquid gates is to prevent thickness variations in the transparency or filter from distorting wavefronts. This is accomplished by immersing the transparency or filter in a liquid of the same refractive index. The liquid is confined by glass plates worked to a high degree of flatness. Only the outer surfaces of the glass need be worked, as the liquid will also partially compensate for variations on the inner surfaces. When a glass plate is used for the filter or object, then a simplified liquid gate may be possible. The need for a liquid gate is seen to depend on the severity of thickness variations and the severity of their effects in the filtering scheme chosen.

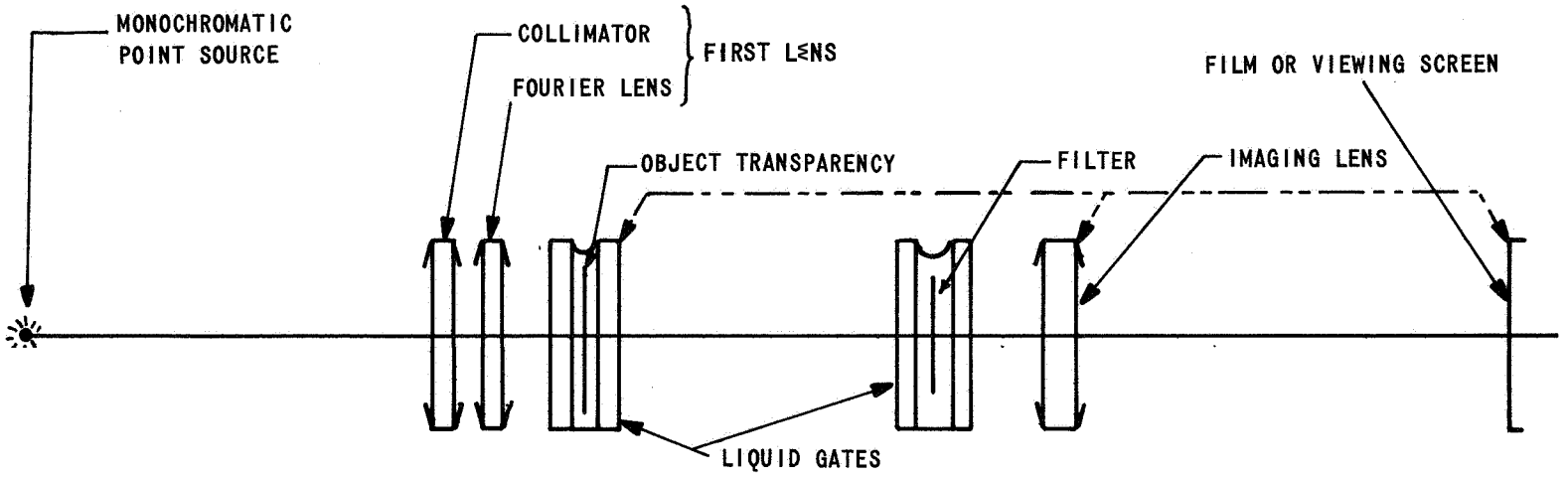


Figure 2 (P2) SPATIAL FILTERING APPARATUS

The first lens shown is often composed of two telescope achromats placed with infinite focus sides together; this is done because telescope achromats are readily available and the combination is corrected for such use, except as noted below.

When this is done, the first lens is called the collimator, and the second lens is called the Fourier lens. If it is desired to process at fixed scale, the object transparency may be inserted between the collimator and the Fourier lens. This layout permits one to use, or not use, a liquid gate without changing the spherical aberration of the Fourier lens and liquid gate combination. To secure variable scale, the transparency may be placed behind the first lens and scale changed by motion toward the filter. In this case, the combination of two telescope objectives is properly corrected for operation without a liquid gate, and a different lens is required for operation with a liquid gate. Fortunately, the spherical aberration is not changed as the gate and transparency are moved to obtain change in scale (see Ref. 5, P. 35). The imaging lens is required to image the object transparency on the receiving film or viewing screen. If the scale of processing is to be varied, the object transparency, imaging lens, and viewing screen are moved together, as shown. The imaging lens should, in most cases, be a specially designed lens to compensate for the changes in spherical aberration caused by the liquid gates if used, the change in field curvature caused by placing the object on the opposite side of the Fourier lens if this is done, and to correct for any aberrations introduced by the filter. Offner (Ref. 6) discusses aberrations in holographic systems.

The principle of operation is that, except for a quadratic phase term later compensated by the imaging lens, the complex amplitude of light in the filter plane, which is conjugate to the source, is proportional to the two-dimensional Fourier transform of the complex amplitude transmittance of the object transparency. This statement may be derived from the Huygens-Fresnel principle if it is assumed that the transparency is uniformly illuminated and that all angles involved are small. The complex amplitude transmittance will be the square root of the ordinary or energy

transmittance for a transparency in a liquid gate. The complex amplitude transmittance of the filter is the complex conjugate of the two-dimensional Fourier transform of the desired reference function. In passing through the filter, the complex amplitude is multiplied by the transmittance of the filter; the correlation is effected in accordance with the convolution theorem for Fourier transforms. The resulting wave continues through the system and forms the filtered image. However, film or the eye is an energy detector. Thus, the overall process takes the square root, correlates and squares. So long as the resulting correlation is positive, this is not a serious problem because contrast may be adjusted by choice of film, developer, time, temperature, etc,

Corresponding to the negative light difficulty in cross-correlators, spatial filtering has the problem of phase in the spatial filter. The phase problem is a practical and not a fundamental problem. The phase considered is the phase in the p olar representation of the complex amplitude transmittance, and corresponds to the optical path length, multiplied by $\frac{2\pi}{\lambda}$ where λ = wavelength. Because one may generate small relative optical path length differences from large relief distances by making the relief surface one between two media of nearly the same refractive index, and because the scale in the filter plane is proportional to the distance from the transparency, it is fundamentally possible to fabricate the phase part of the filter in a milling machine. Investigators and inventors have had some success in generating these phase variations at reasonable scale, and in simulating them.

The simplest phase problem is that of fabricating or simulating phases of only 0 or π; i. e., the optical path lengths at all points of the filter are all in one of two classes, all of those in the first case are to be the same, and all of those in the other are to differ from those in the first only by one-half wavelength. This is an important special case, as it corresponds to the removal of any defect with rotational symmetry.

Such phase differences may be fabricated by the use of processes in which the thickness of a photographic emulsion is controlled by effects taking place when the emulsion is developed (the silver is bleached out after development). Also, a method of simulating phase changes of π is discussed in Reference 7.

The only proven method of simulating arbitrary phase is the spatial carrier method, discussed in Reference 8, Chapter V. The function supplied is not the desired filter function but rather $|F(\nu_x,\nu_y) + B\exp(-iC\nu_x)|^2$ where

$F(\nu_x,\nu_y)$	=	desired filter function
B, C	=	constants

The expansion of the above expression is

$$|F(\nu_x,\nu_y) + B\exp(-iC\nu_x)|^2 = |F|^2 + |B|^2 + FB^*\exp(iC\nu_x) + F^*B\exp(-iC\nu_x)$$

The combination $FB^*\exp(iC\nu_x)$ is equivalent to F followed by a prism. Thus, the filtered image appears off axis. The combinations $|F|^2 + |B|^2$ and $F^*B\exp(-iC\nu_x)$ generate images filtered by different functions, but they appear on axis and displaced to the other side of the axis, respectively. This method bears a close relationship to holography, or imaging by wavefront reconstruction so that spatial filtering with a carrier is sometimes called holographic image restoration. In holography, the filter function is the transform of the desired final image. The final image is obtained by illumination of the filter function, or hologram, with a plane wave, which in the formulation above is equivalent to convolution with a delta-function in both filtering and holography, it is possible to replace the term $F(\nu_x,\nu_y)$ by a term which includes a quadratic phase. Such replacement gives the hologram (or filter) focusing properties, and is called a change from Fourier to Fresnel holography. This replacement is important in holography as it makes possible "lensless photography" but use of Fresnel transforms is of no advantage in spatial filtering.

An interesting possibility, however, is that for fixed scale operation, it may be possible to use the filter to correct spherical aberration

in the imaging lens. Because coma and distortion may be eliminated by the use of symmetry, and chromatic aberrations are of no effect in a monochromatic system, the full efforts of lens design could be brought to bear on astigmatism and field curvature, resulting in a near-perfect lens. This would contrast with the usual problems of design of imaging lenses, where it is necessary to compromise between spherical aberrations and astigmatism, (See Ref. 5, especially Chapter XIX).

As noted above, the spatial carrier method is the only proven one if arbitrary phase is required. If phase is not required, the spatial carrier method should not be used, and if only simple phase changes of 0 and π are required, other methods should be considered, because of the following disadvantages:

1) The spatial carrier frequency must be high enough to separate the desired and undesired beams, this results in some combination of high resolution film or plate that the filter is on, inconveniently large apparatus, and/or small format for the object transparency.

2) Light lost by being diffracted into unused beams.

3) The off-axis operation aggravates mechanical problems in fabrication and alignment.

4) The spatial carrier introduces unique aberrations not found in centered systems; these may, however, be corrected by the use of a prism as discussed in Reference 7 at the cost of added complexity.

3. Detail Processing by Analog Electronics

The method of operation of analog image processors is to convert the image to an electrical time function, operate on this function by electrical filtering methods, and to display the processed image by use of a Kinescope, which may be photographed for a permanent record. Kovasnay (Ref. 9) discussed such a processor, using a flying spot scanner.

The basic difficulty of this type of processing is that the images are two-dimensional, but the analog electronics are one-dimensional. Thus, the suitability of such processes depends on the complexity of the reference function, in the sense that a more complex reference function requires more one-dimensional functions to simulate it. A two-dimensional reference function can be built up as a sum of one-dimensional reference functions, each corresponding to a scan of the two-dimensional function and the total cross-correlation is then the sum of the cross-correlation with the scans. But, it is necessary to add a bias to each of the partial correlations to prevent loss of information should the partial correlation become negative. If the reference function is complicated, then it is expected that the total bias, which is the sum of all the biases, will be so large as to swamp the signal. Therefore, if the reference function is at all complicated, then this method is inferior to the method discussed previously of forming correlations first and then scanning.

4. Detail Processing by Digital Computation

The digital computer method of image processing can be very powerful but generally requires long computing times and is, therefore, expensive. The limits on reference functions, inherent to modern digital computers, are beyond imaginable needs of image processing. Computer techniques provide a method of quasi-homogeneous processing. Also, as the services of these computers may be hired, it may be considered that the computer takes up no space, regardless of the size and complexity of the image to be processed. Computers also have disadvantages in the preparation of images for digital processing and in recovery of continuous tone imagery from the digital output.

For digital computer processing, the image and the reference function are supplied as discrete samples, usually spaced equally in two dimensions. The validity of the replacement is based on the sampling

theorem as discussed previously and in Ref. 1 and 2. The correlation integral becomes a sum, and the required operations are only addition and multiplication. The precision of modern digital computers is such that round-off errors are negligible with respect to image processing. The sample values of the reference function, or weights, are constrained to certain ranges, but these ranges are large compared to reasonable ranges of desired weight values, and may be negative. Hence, as noted above, there are practically no limits on reference functions. By changing the weights during the processing, quasi-homogeneous processing may be achieved. Furthermore, there are no theoretical limits to the complexity of the weight-changing scheme.

One disadvantage of digital processing is that it should, in many cases, be both preceded and followed by a halftone stripping type of process. The requirement for stripping an input is that although there is no signal at spatial frequencies beyond the diffraction limit of the taking lens, there usually is noise there. Also, it may not be considered worthwhile to try to save the signal at these high frequencies, because to do so would require a greater sampling rate. Thus, it is necessary to remove components at these high frequencies lest they be folded into the signal. The requirements of this processing should completely remove modulation frequencies above the cutoff without disturbing modulation at frequencies below the cutoff. For analog automatic stereo compilers, the halftoning must be removed to avoid spurious effects, as discussed in Appendix P-3, and for human stereo compilers, or stereointerpretation, it must be removed to avoid bothersome Moiré fringes. As discussed later, spatial filtering offers the best performance in halftone stripping.

The reader can easily visualize situations where a three-step process is required: 1) prepare the image for digital processing by spatial filtering, 2) process digitally, 3) recover the continuous tone image by spatial filtering. Clearly, a considerable simplification would result if the processing could be done by one spatial filtering process, which is the

case if truly homogeneous processing is required. Because truly homogeneous restoration for automatic stereo compilation is of no value, see Appendix P-3, and it has not been discovered how to do quasi-homogeneous processing by spatial filtering, the three-step process is required for automatic compilation.

A digital computer is usually used for inhomogeneous processing, as the operations are too complicated to permit simple analog electrical or optical implementation. The difficulties of using sampled data, mentioned above for homogeneous processing, apply to inhomogeneous processing as well. Also, inhomogeneous schemes should be checked to determine if round-off errors can cause problems.

5. Detail Processing to Remove Spurious Structure

Removal of halftone and line scan structure is a special case of homogeneous processing. It can be performed by incoherent cross-correlators, by spatial filtering, or by special incoherent halftone strippers. These incoherent strippers are discussed later and the effects of their use are compared to those of spatial filtering, and defocus of the enlarger which is a non-optimum case of optical cross-correlation. Most of the discussion deals with removal of halftone.

Incoherent strippers were developed for use in graphic arts to prevent Moiré fringes in the rescreening of halftones. That is, sometimes it is necessary to make a halftone of an image which is already a halftone. When this is done, the old halftone structure must be completely removed, because if any of it were allowed to remain, it would interfere with the new halftoning and so create annoying effects.

As discussed later in this appendix, an ideal stripping process would have unit response over the region of the spatial frequency domain where there is modulation at frequencies proportional to those in the original object and would fall to zero upon leaving the region and remain at zero for all frequencies where there is modulation at frequencies which are sums and differences of scaled object frequencies and halftone harmonics.

Operation of the incoherent strippers is based on the diffraction theory of aberrations (see Ref. 10) which states that the frequency response of the lens in the camera used to copy the halftone is the autocorrelation of the pupil function, suitably scaled. The stripper is a glass slide covered with spots in some pseudo-random pattern. The spots are sufficiently thick to generate one-half wavelength of optical path difference relative to air at some nominal wavelength and cover half of the area. Thus, they introduce into the pupil function of the lens the same group of spots. It may be seen that for large spatial frequencies, corresponding to large shifts of the pupil function for autocorrelation, there will be no correlation and hence, no response. Thus, they satisfy the requirement of rejection of information at the sum and difference frequencies, But they do not satisfy the requirement for full response up to the cutoff frequency, as the response of such a system must start to fall immediately on leaving zero. The same general features would have been obtained by stopping down the camera lens so that its diffraction limit would fall at the cutoff frequency, but this would result in an inconveniently long exposure time.

The frequency response of a defocused camera is given in Ref. 11 for small values of defocus. For large values, one may assume geometrical optics, and derive that it is:

$$\mathcal{T}(\omega) = \frac{J_1(K\omega_r)}{K\omega_r} \qquad \text{(P2-1)}$$

where

$J_1(\ )$ = Bessel function, first kind, first order

ω_r = Radial coordinate in spatial frequency plane

K = Constant (depends on geometry)

In all cases, the defocus response is seen to be an oscillatory function of frequency. Thus, defocus cannot be expected to completely remove undesired high frequencies. Nevertheless, defocus is often used in re-screening because the defocused response is greater in the region of desired frequencies (for equal values of first zero of response).

It is theoretically possible to remove halftone with a correlator, to the specification outlined above. The filter would require negative parts.

Spatial filtering is the best method of removing halftone from the standpoint of image rejection and salvage of detail. A blackened piece of metal with a hole in it placed in the filter plane satisfies the requirement of full response out to the cutoff, and zero beyond. This filter does not require a liquid gate. **Also,** in most cases, the liquid gate for the object transparency is not expected to be required. The disadvantages are in convenience. First, a spatial filtering apparatus is required. Second, it is necessary to have the image to be processed in the form of a transparency. Third, the exposure time is likely to be longer than with the defocus or special incoherent process unless a laser is used as a light source for the spatial filtering. The advantages of the spatial filtering will outweigh the disadvantages where the best performance is required. A comparison of the halftone removal methods is given in Table I (P2) below.

TABLE I (P2)

CHARACTERISTICS OF HALFTONE STRIPPING PROCESSES

	Performance in Pass Band	Performance in Stop Band	Convenience
Special Incoherent	Poor	Excellent	Excellent
Stopping Lens	Poor	Excellent	Poor
Defocus	Fair	Fair	Very Good
Spatial Filtering	Excellent	Excellent	Fair
Cross Correlation*	Very Good	Very Good	Poor-Fair*

* Pending solution of negative light difficulty

6. Detail Processing to Remove Modulation at Very Low Spatial Frequencies

Image processing machines which primarily suppress modulation at very low spatial frequencies are called dodgers. The suppression of low frequencies occurs because the transmittance of the negative at a point is weakened by the average of the transmittance over some surrounding area. The machines are limited by their type of construction to this type of processing. Automatic dodging machines are commercially available.

IV. DETERMINATION OF DATA USED IN MAKING CORRECTIONS AND OF CORRECTIONS TO BE APPLIED

An important part of any image processing scheme is the determination of the degradations. For tone and distortion, the restoring operations are simply the inverse of the degrading operations, so that when the degradations can be measured, the restoration is straightforward, The perfect restoration of detail is not possible in the presence of noise. For detail processing there are two problems, first to measure the degradations, and second to decide which, if any, processing will result in an improved image.

A) Determination of Correction Data for Tone

The theory of determining the corrections for tone restoration is straight forward; the response of the film or other sensor is to be determined by supplying a series of known control input radiance levels and recording the outputs, Then, the corresponding input for any output may be found by interpolation. If this is done graphically, with equal steps of log input, the familiar H & D curve is used (Ref. 14). The inputs are to be supplied over an area sufficiently large to prevent degradation of the results by noise and the detail rendition properties of the sensor. The use of large constant input areas results in availability of only a few input levels, but as the function to be determined is one-dimensional and the inputs may be supplied equally spaced, there is no fundamental difficulty in interpolation and smoothing.

There appears to be no practical method to implement the above scheme in space photography to perfection, but useful correlation data can be obtained. For a complete simulation, the control exposures should be given the film at about the same time as the image exposure, both should last for about the same length of time, and should use the same distribution of energy over wavelength (Ref. 15). The usual methods of supplying the control exposures are 1) use of control panels of known reflectance, 2) giving parts of film known exposure in a special machine known as a sensitometer shortly before takeoff and/or after landing, and 3) photographing standard light sources shortly before takeoff, etc. It should be noted that for the last two methods, care must be exercised to use the same wavelength distribution. Also, the last two methods require separate corrections for atmospheric radiance, and the second, for flare. The first method does not appear possible, because of the difficulty of transporting standard panels into space, and the uncertainty whether their reflectances would change. The other two methods can be helpful but cannot yield the best results because film is perishable and its properties may change during the long time between image and control exposures. Including a sensitometer in a spacecraft to provide control exposures is considered a poor idea because of the space, weight, and perhaps power required. Fortunately, as discussed previously, the requirements for tone restoration are not severe so that the use of control exposures supplied prior to launch is expected to be adequate.

B) Determination of Correction Data for Distortion

The determination of the distortion data is part of camera calibration (see Ref. 3, p. 173ff). In measuring distortion, one has the choice of using the camera as a projector, supplying targets at known format positions in image space, and measuring the corresponding angles in object space or of using the camera as a camera and measuring the

locations of images of targets at known angles. For the camera method, the targets may be artificial stars supplied by a laboratory instrument called a multicollimator, an array of targets on the ground photographed from a tall tower, or the natural stars. The last mentioned method would appear to be most advantageous for space topography as it permits calibration of the camera in its natural weightless condition, and after the shock of launch. The disadvantage of natural star calibration appears at the data reduction stage. The distortion is available at the star images, and an interpolation scheme is required to find the distortion at all points. Also, as there are likely to be errors in the calibration itself, it would be desirable to fit a distortion function with considerably fewer degrees of freedom than the number of stars visible; i.e., to do some smoothing. Various mathematical methods are available to interpolate and smooth in some sense equally spaced data; the natural stars, however, are where one finds them. Further mathematical research is required to find and evaluate mathematical techniques for fitting arbitrarily located data consistent with the calibration problem.

C) Determination of Correction Data and Operations for Detail Processing

1. Determinations for Homogeneous and Quasihomogeneous Processing

The determination of a proper reference function or spatial filter for detail processing requires two steps. First, the point spread or its two-dimensional transform, the modulation transfer function, must be found. If noise with a spectral distribution which is not constant over the region of interest is involved, then the power spectrum of the noise is required also. Then, on the basis of the above data, the reference function or spatial filter must be designed.

a) Measurement of Point Spread - The accurate determination of point spreads of imagery taken for use (as to be distinguished from those taken for calibration) is very difficult because they are greatly

affected by operational deficiencies. Therefore, it would be advisable to measure the point spread of any camera at several focal settings before use as an aid in determination of the operational point spread. These point spreads are composed of contributions from lens aberrations, lens defects, defocus, uncompensated image motion, and the point spread of the sensor involved. The sensor point spread may validly be combined with the motion point spread by convolution, but the effects of aberrations, defects, and defocus must be combined as a whole by other methods before being convolved with results of the previous convolution to arrive at the overall point spread. That is, defocus cannot be considered separately from aberrations and defects. As defocus and uncompensated motion cannot be predicted in advance, (if they could have been predicted, they would not have been permitted to exist), the overall point spread cannot be predicted in advance. Thus, at least partially, the point spread must be determined by an analysis of the imagery or by use of instrumentation not usually included in photographic missions. But, prior measurement of the film spread function and of the lens spread function at several focal settings would be helpful. If such were done, the imagery analysis need yield only six numbers*, giving the amount and direction of uncompensated motion and of defocus rather than a two-dimensional function. As the analysis of imagery taken for use is made inaccurate by noise and limited availability of sub-images suitable for use as test targets, it is wise to require as little as possible from this analysis so that maximum redundancy will be available.

The determination of the point spread or MTF of the 1ens or the film may be performed by analysis of images of special targets. Favorites among these are points and one-dimensional sine waves. The determination of the MTF of lenses may also be done by use of special interferometers. The determination of the transfer function of the entire system is usually based on the analysis of edges.

---

* Two linear velocity components, three angular velocity components, and the defocus.

b) Choice of Reference Functions

Appendix **P-4** contains a set of schemes for arriving at proper restoring filters for automatic stereocompilation. These may also apply to manual compilation; it is recommended that this possibility be investigated. For processing of data for interpretation, no comprehensive theory is available, and one must operate on the basis of general rules and considerations. To arrive at the modulation transfer function (MTF) of a restoring filter for interpretation, it is sometimes helpful to consider the MTF of the system used to form the degraded image, and from this to arrive at the MTF of the overall system. Then the MTF of the restoring filter may be found by comparison of the above two. It is to be noted that if the degradation MTF has a zero, then the overall MTF must have a zero at the same place, hence this detail processing is not the taking of a simple inverse, as is the case for tone and distortion. The following are some considerations for the overall MTF:

1) So long as there is adequate DC to prevent images from becoming negative, there is no need for modulation at very low spatial frequency, corresponding to variations apparent only over areas larger than those to be imaged on the fovea during interpretation. Therefore, these may as well be eliminated permitting use of higher contrast.

2) There is no need for the restoring filter to transmit noise at spatial frequencies so great that no signal is present.

3) If the MTF of the system not including processing contains negative regions, these should be rectified, or at least discarded, to eliminate spurious resolution.

4) To minimize ringing,* the composite MTF surface should be continuous and perhaps even smooth everywhere there is signal.

---

* Ringing occurs when the point spread is an oscillatory function of distance. This corresponds to discontinuities in the MTF or its derivatives in regions where signal modulation exists. It causes the images of small, bright objects to be surrounded by one or more outlines or rings of alternating contrast.

The tolerance on ringing is not known. Some is obviously permissible, as a diffraction limited lens is often considered the model of perfection, yet has a ringing point spread. But it is known that severe ringing is objectionable. This means that it will be necessary to reduce the MTF near zero crossings so that it may approach zero smoothly.

5) The statements in 3) and 4) above are based on the assumption that the degraded image is to be converted into something approaching a normal exposure. There is also the possibility of conversion to a double exposure. If the double exposure goal is selected, then (4) is inapplicable and (3) is to be interpreted as to assign the proper sign to the restoring filter and be consistent with the double exposure, rather than to make all regions positive.

Because the methods of inhomogeneous image processing are not determined in all cases, it is not possible to specify the required inputs and operations in all cases. Consider first the inverse imaging approach to inhomogeneous processing. This method ignores noise and attempts to invert the imaging operation, expressed as a convolution, by the solution of the resulting integral equation of the first kind. The point spread appears as the kernel in this approach. To avoid singularity of the integral equation, certain boundary conditions are required. It is apparent that the size of the area over which these conditions must be specified is related to the nature of the point spread because, in one dimension, the diffraction limited point spread which extends to infinity requires knowledge of the object from fixed points to positive and negative infinity, but restoration of uniform image motion (considered as a convolution with a flat-topped, vertical-sided pulse) requires knowledge of the object only over an interval of the same length as the pulse. In summary, inverse imaging requires knowledge of the point spread and of boundary conditions consistent with the point spread.

The operations used to perform inverse imaging are, of course, the operations used to solve integral equations; viz. 1) numerical methods and 2) expansion in orthogonal eigenfunctions. The numerical methods are straightforward, the degraded image and the point spread are replaced by discrete data sets. The imaging convolution integral becomes a weighted summation. The radiance at each point in the degraded image may be expressed as a weighted summation of object radiances, i. e., it generates a linear equation. Assuming adequate boundary conditions, the set of all linear equations corresponding to all image radiances may (at least theoretically) be solved to obtain the object. Unfortunately, this often can not be done for reasonable object sizes because of round-off errors. There are some good examples of this in Reference 16 and a method of avoiding these effects by smoothing is exhibited.

An example of the other method of attack, expression in orthogonal eigenfunctions, is given in Ref. 13. If these functions are ordered with the least oscillatory first, as is customary in ordering orthogonal functions, then it may be hoped that a good solution may be obtained by truncating the expansion at the proper function. Because the actual expansion will involve numerical integrations, the possibility of round-off errors still exists. Also, both of the above methods are likely to be very sensitive to noise in the degraded image.

Because noise is considered as sort of an after thought in the inverse imaging attack, it is unlikely that it results in an optimum processed image. A rigorous inhomogeneous processing attack would consider: 1) the point spread, 2) the boundary conditions, 3) the characteristics of the noise, and 4) the criterion of performance. A criterion is necessary because it was necessary to give up perfect restoration to combat noise. It is submitted that a more fruitful method of attack could be through statistical estimation theory (Ref. 17, 18) as it considers (1), (3), and (4) above and some progress has been made on (2), but inverse imaging considers only (1) and (2). Because statistical estimation theory was developed in

connection with multipath transmission of radio signals, some translation is necessary to convert it to image processing. This is given in the paragraph below.

In the analysis of multipath transmission, a known function is transmitted, it is distorted by the multipath, noise is unavoidably added, and it is desired to estimate the characteristics of the multipath considered as a linear filter by use of a linear estimating filter as shown in Figure 3a (P2). The estimating filter may be time-varying (=inhomogeneous). The problem is formalized, as shown in Figure 3b (P2). The known signal s (t) is applied to the unknown filter with impulse response D (t) to the output of which is added the noise sample N (t). It is desired that the output of the estimating filter shall approximate the impulse response of the unknown filter. This is equivalent to the system shown in Figure 3c (P2) where the known function is generated by application of an impulse to a known filter. And, this is equivalent to the system in Figure 3d (P2) where the order of the filters has been interchanged. The impulse generator and unknown filter may be taken together as the generator of an unknown function, shown in Figure 3e (P2). Finally, this may be (after generalization to two dimensions) interpreted with respect to image restoration as shown in Figure 3f (P2). Hence, it is expected that the results of these theories may be carried over and used as a base for inhomogeneous image restoration.

The requirements for a halftone removal filter are best explained in the spatial frequency domain. Recall that multiplication in the spatial domain corresponds to convolution in the spatial frequency domain, and conversely. A halftone of a uniform gray area consists of a two-dimensional Dirac comb convolved with the dot shape. Thus, the transform of a halftone is that shown in Figure 4a (P2) consisting of the product of a two-dimensional Dirac comb, which is the transform of itself, and the transform of an individual dot. The spectrum of an actual optical image is shown in Figure 4b (P2) where regions with non-zero modulation are shown shaded. The

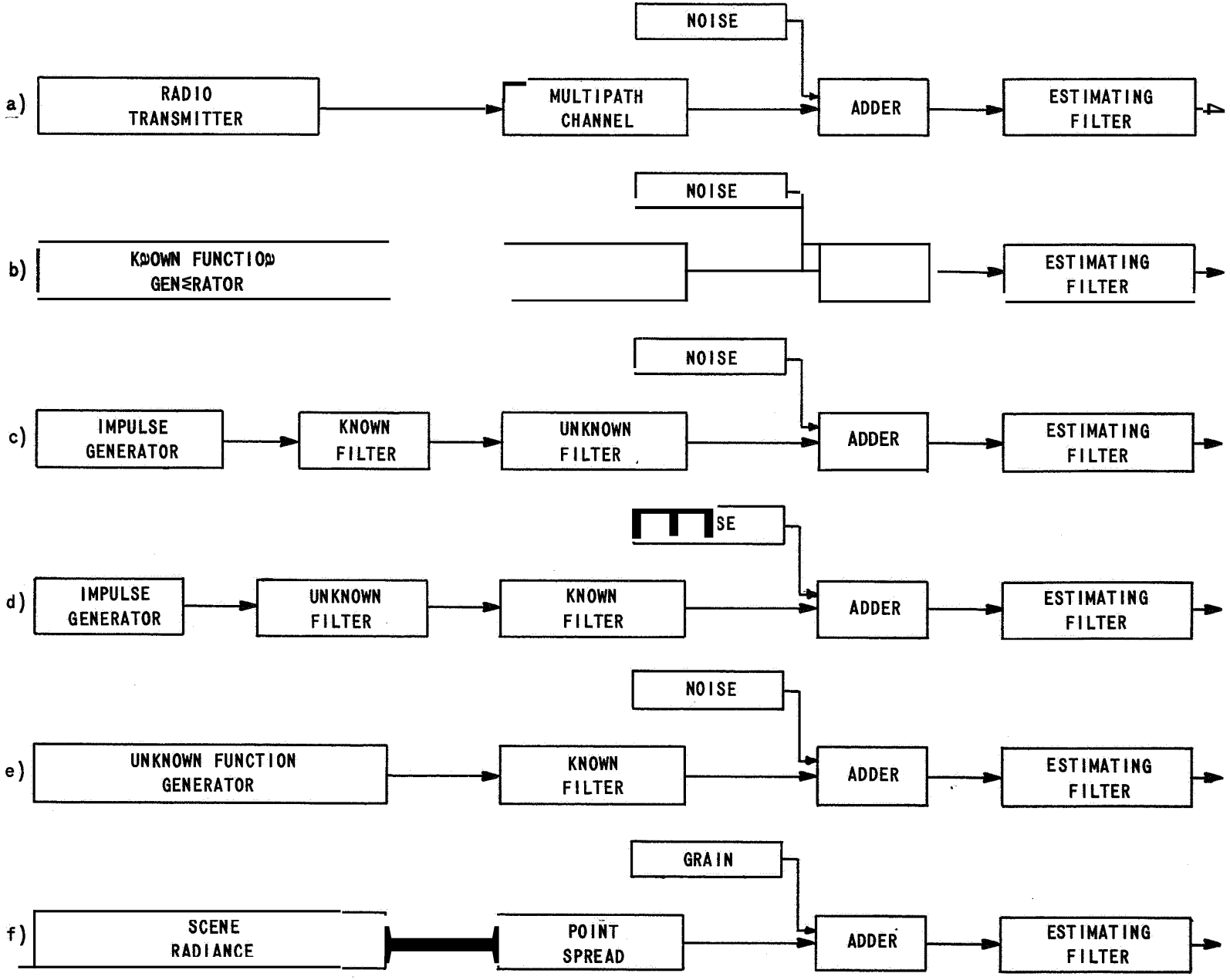


Figure 3 (P2) ILLUSTRATING CONNECTION BETWEEN MULTIPATH TRANSMISSION ANALYSIS AND INHOMOGENEOUS IMAGE RESTORATION

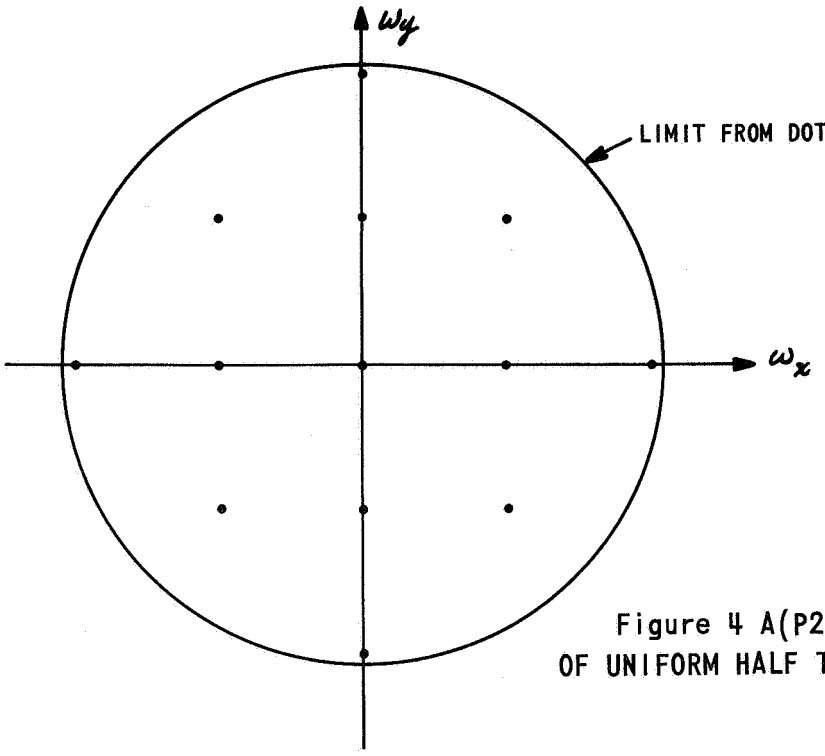


Figure 4 A(P2) SPECTRUM OF UNIFORM HALF TONE EXPOSURE

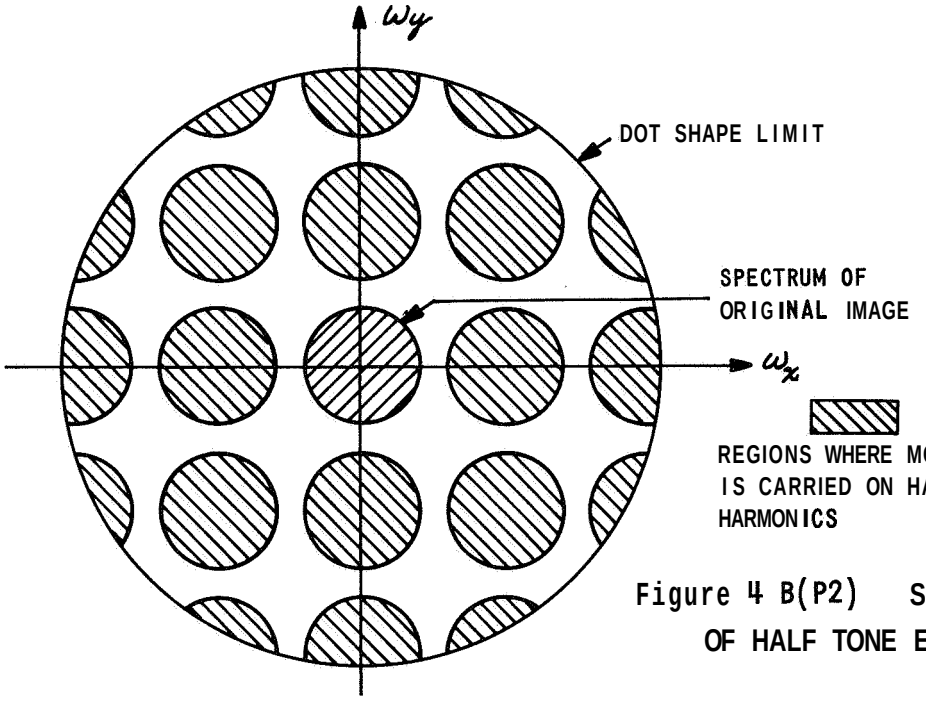


Figure 4 B(P2) SPECTRUM OF HALF TONE EXPOSURE

halftone exposure is the product of a uniform grey halftone and the original image, denoted by ////, and the other components are spread are also denoted by \\\\. The spectrum of the developed halftone will differ from that of the exposure because of the non-linear content transfer of the photographic process, but this does not change the main points of the argument. It may be seen that to thoroughly suppress the halftone structure, one must block all of these high scanned difference frequencies. Also, it is desirable to pass all of the main image frequencies without attenuation so that freedom will exist to progress as discussed previously. The above applied to images which were screened with a screen fine enough to move the modulation carried on the halftone harmonics enough away from the main spectrum to prevent overlap. When overlap occurs, it is not clear what the best procedure is.

The exact shape of the point spread for dodging is not critical. Thus, it is permitted to take whatever shape is convenient to simplify mechanization, consistent with the requirement that it consist of a small, positive area surrounded by a large, negative one.

V. CONCLUSIONS AND RECOMMENDATIONS

The following conclusions and recommendations concerning image processing for space topography are drawn from the previous discussion.

1) Expend no research effort on restoration of tone or grey scale.

2) Investigate the mechanism of destructive scan readout distortion as an aid to the search for possible restoration schemes.

3) Perform correction of distortion by mechanical or numerical methods.

4) Conduct mathematical research on surface fitting with arbitrary data point locations.

5) If detail restoration is contemplated, determine MTF of film and of lens at several focal settings.

6) Conduct research to more clearly define the objectives of homogeneous image processing for the interpretation of detail.

7) Determine if surfaces to be mapped have uniform color, like the moon, at earliest opportunity, to determine reasonableness of processing for off-axis aberrations.

8) Attempt to discover methods of quasi-homogeneous processing which do not require sampling.

9) Conduct research to determine if conclusions in Appendix P4 regarding automatic compilation are valid for manual plotting also.

10) Conduct research in inhomogeneous processing.

11) Perform line-scan and halftone stripping by coherent spatial filtering when best performance is required.

VI. REFERENCES

1. Shannon, C. E. and Weaver, W., The Mathematical Theory of Communication, The University of Illinois Press, Urbana, Illinois, 1949, p. 53.

2. O'Neill, E. L., Introduction to Statistical Optics, Reading, Mass., Addison-Wesley, 1963, p. 163.

3. Thompson, M. M. (Ed.), Manual of Photogrammetry, Falls Church, Va., American Society of Photogrammetry, 1966.

4. Trabka, E. A., and Roetling, P. G., "Image Transformation for Pattern Recognition Using Incoherent Illumination and Bipolar Aperture Masks", J. Opt. Soc. Am., Vol. 54, 10, October 1964, pp. 1241-1252.

5. Conrady, A. E., Applied Optics and Optical Design, New York, Dover, 1957.

6. Offner, A., "Ray Tracing Through a Holographic System", J. Opt. Soc. Am., Vol. 56, 11, November 1966, p. 1509.

7. Holladay, T. H., and Gallatin, J. D., "Phase Control by Polarization in Coherent Spatial Filtering", J. Opt. Soc. Am., Vol. 56, 7, July 1966, p. 859.

8. Stroke, G. W., "An Introduction to Coherent Optics and Holography", New York, Academic, 1966.

9. Kovasznay, L. S. G., and Joseph, H. M., "Image Processing", Proceedings IRE, Vol. 43, May 1955, pp. 560-570.

10. Born, M., and Wolf, E., Principles of Optics, Second Edition, New York, Pergamon, 1964, pp. 480-490.

11. Hopkins, H. H., "The Frequency Response of a Defocused Optical System", Proceedings of the Royal Society of London, Series A, Vol. 231, pp. 91ff, 1955.

12. Harris, J. L., Sr., "Image Evaluation and Restoration", J. Ope. Soc. Am., Vol. 56, 5, May 1966, p. 569ff.

13. Barnes, C. W., "Object Restoration in a Diffraction Limited Imaging System", J. Opt. Soc. Am., Vol. 56, 5, May 1966, p. 975ff.

14. Mees, C. E. K., Theory of the Photographic Process, New York, MacMillan, 1954, pp. 852-854.

15. Gorokhovskii, Y. N., Spectral Studies of the Photographic Process, translated from the Russian by Grace E. Lockie, New York, Focal 1960, pp. 83-94 and 724-727.

16. Phillips, D. L., "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind", J. Association for Computing Machinery, 9 (1962)- pp. 84-97.

17. Levin, M. J., "Optimum Estimation of Impulse Response in the Presence of Noise", 1959 IRE Convention Record, Part 4, pp. 174-181.

18. Turin, G. L., "On the Estimation in the Presence of Noise of the Impulse Response of a Random, Linear Filter", IRE Trans. on Information Theory, Vol IT-3, pp. 5-10, March 1957.

19. "A Study of Automatic Techniques for the Selection and Identification of Targets (U)", CAL Report No. VG-1542-G-11, Technical Documentary Report No. AL-TDR-64-164, September 1964, pp. 19-27, CONFIDENTIAL.

20. Muerle, J. (Ed.), "A Study of Automatic Techniques for The Selection and Identification of Targets (U)", Technical Report AFAL-TR-66-357, Air Force Avionics Laboratory, WPAFB, Ohio, December 1966, pp. 119-147, CONFIDENTIAL.

VII. BIBLIOGRAPHY

1. O'Neill, E. L. "Spatial Filtering in Optics", IRE Trans. on Information Theory, IT-2, (1956), P. 56ff.

2. Schade, O. H. , "Optical Image Evaluation", Nat. Bur. Stand. Circ., 526, p.231ff, 1954.

3. Cutrona, L. J. et al, "Optical Data Processing and Filtering Systems", IRE Trans. on Information Theory, IT-6, June 1960.

4. Roetling, P. G., and Hammill, H. B., "Study of Spatial Filtering by Optical Diffraction for Pattern Recognition", CAL Report No. VE-1522-91, FEB 62.

APPENDIX P-3

IMAGE RESTORATION FOR AUTOMATIC STEREOCOMPILATION

By John D. Gallatin

I. INTRODUCTION AND SUMMARY

This appendix derives suitable processing filters to be used on imagery prior to automatic stereocompilation. The discussion is limited to that part of map compilation which consists of finding spot altitudes and contour lines. The derivation is based on automated compilation, but may be applicable to manual compilation. Other applications of image processing filters are discussed elsewhere in this report.

From the investigations discussed in this appendix, it is concluded that if both images are degraded by the same modulation transfer functions (MTF), no irnage restoration is required or should be attempted, and that if they have been degraded by different MTF's, the optimum filtering consists of giving to each image the degradations of the other, and somewhere using a compromise gain which does not differ greatly in overall effect from the unfiltered case.

Section II contains a brief description of stereocompilation, and especially, automated compilation. Section III contains a description of matched filters and mentions some of their optimality properties relevant to this discussion. In Section III, it is shown how current automated compilers correspond to matched filters if the images are degraded identically. A derivation of image processing filters for differently degraded images is given in Section IV. In Section V, the discussion is extended to situations in which a data link must be considered and the conclusions are assembled and restated in Section VI. The derivations in Sections 111, IV and V are based on maximum likelihood estimation: however, under the assumptions of the discussion this is equivalent to the photogrammetrically traditional mean squared error criterion[1,2].

11. AUTOMATED STEREOCOMPILATION

The essence of altitude determination by stereoscopic parallax is depicted in Figure 1(P3). In the figure, it is assumed that the vehicle's path is parallel to the datum plane and that an ideal camera was used with the axis normal to the datum plane. When these conditions do not hold, suitable corrections may be applied which are not of concern here. As may be seen from the figure, deviations of the surface from the datum plane cause a relative displacement (b) in the rays to the camera stations. From similar triangles, the spot altitude (considered positive down) is

$$h = \frac{b(H+h)}{B} \approx \frac{bH}{B} \qquad \text{(P3-1)}$$

where H = aircraft altitude
and B = base line,

as shown in the figure. Thus, when the orientation and scale of the system as a whole have been determined, the determination of spot altitudes and contours is reduced to determining the parallax (b) at suitable points, or in reality, over small areas.

Although one speaks of spot altitudes of points, these altitudes are based on small images; i. e., small parts of the entire image. Except for a very small, bright source, which is of photogrammetric interest only when it is a surveyed control target, it is not meaningful to consider matching points, as a point is too small to have shape, contrast, texture, etc. Rather, sub-images are matched and the altitude derived from the matching is assigned to some point which belongs to or is near to the portion of the terrain appearing in the sub-images.

The current method of determining this parallax automatically is to form the cross-correlation of the two images and to take as the parallax the abscissa which maximizes the correlation? That is, if $T_1(x,y)$ is the transmittance of the diapositive (a diapositive is a positive image) corresponding to camera station #1, or some monotonic function of it, then the integral

$$R_{12} = \iint T_1(x,y)\, T_2(x+X, y) \qquad \text{(P3-2)}$$

is computed and the value of X which maximizes R_{12} is taken as the parallax. In the above statement, the integration is carried out over a region large compared to the smallest resolvable detail, so that errors are reduced by a sort of averaging process, yet small compared to the region covered by the stereo pair, so that many altitudes may be determined and contours drawn from a pair of diapositives.

Consider, for example, stereocompilation of permanent bleachers at a scale such that the seat numbers may be resolved. Then, by matching images of seat numbers, (assumed painted on the top of bench seats) and of the edges of the seats (these images are the same in both diapositives), one may find a spot altitude which is assigned to all points on the top of that row of seats.

Consider now compilation of a hillside. The detail is supplied by bushes, rocks, etc. The parallax varies within the sub-regions which are being used to measure it, and hence the sub-images are not identical. Thus the correlation may not yield a useful result. If the slope of the terrain (magnitude and direction) were known, then it would be possible to project both images into the datum plane, and from these projections determine the parallax and altitude. Therefore, the sloping terrain problem may be overcome by the use of a suitable iterative procedure. This procedure would use previously determined altitudes to estimate slopes, These slopes may be used for local projection so that more refined values of altitude may be determined. The projection may be accomplished with zoom cylindrical cases.

In the above discussion, as noted in the figure, the x direction is the direction of motion from station #1 to station #2. The integral may be reduced to a single integral by scanning in the x direction, this is a common mechanization. Also, rather than compute the entire correlation function, it is sometimes computed for *two* values of X, hopefully one on each side of the maximum, and a control system is used to equalize these values by changing the X values. Then the parallax is taken as being midway between the X values.

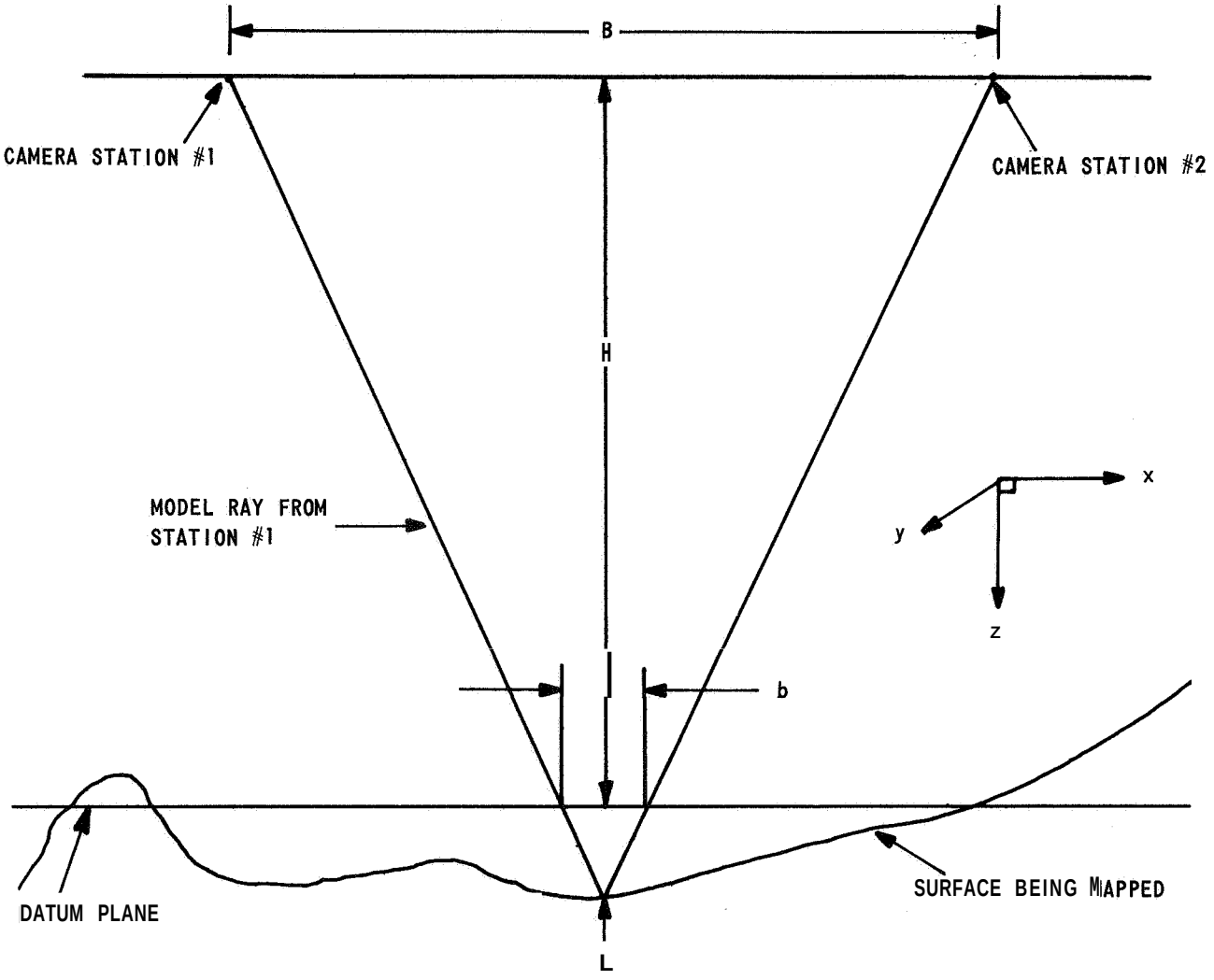


Figure 1 (P3) GEOMETRY OF STEREO PROJECTION

III. MATCHED FILTERS

A matched filter for white noise, in the electrical sense, is one whose impulse response is the time reverse (with a delay) of the signal to which it is matched. The time is reversed because electrical filtering operations are expressed as convolutions. If one considers filtering operations as correlations, then the white noise matched filter is one whose impulse response, or line spread, is the same as the signal. "White", above means that the noise has, in the limit of an infinite sample, equal power at all frequencies of interest.

The generalization of the matched filter to spectrally dependent noise is based on Fourier transforms. If it is sufficiently well behaved, a function of x may be expressed as

$$f(x) = \int_{-\infty}^{\infty} F(\nu) \exp(2\pi i \nu x)\, d\nu \qquad \text{(P3-3)}$$

where $F(\nu)$ is computed from,

$$F(\nu) = \int_{-\infty}^{\infty} f(x) \exp(-2\pi i x \nu)\, dx \qquad \text{(P3-4)}$$

In this representation, the white noise matched filter is the one whose frequency response, or transform of impulse response, is the complex conjugate of the transform of the desired signal. For spectrally dependent noise, the matched filter is the one with frequency response

$$M(\nu) = \frac{S^*(\nu)}{|N(\nu)|^2} \qquad \text{(P3-5)}$$

where $S(\nu)$ = transform of signal

$|N(\nu)|^2$ = quadratic spectrum of noise suitably averaged

$*$ = denotes complex conjugate.

Either of these filters may also contain a phase term linear with frequency which corresponds to a fixed time delay.

An important property of matched filters, in fact, the one which led to their inclusion in this appendix is that they may be used to obtain maximum likelihood estimates. A maximum likelihood estimate of any parameter is obtained by choosing that parameter value for which the probability that its corresponding signal could have resulted in the observed signal corrupted by noise is maximum. If a known signal is delayed and corrupted by additive, gaussian noise, the maximum likelihood estimate of its delay, or parallax, may be obtained by passing the signal into a matched filter and choosing for the delay, or parallax, the time (position) when (where) the output is a maximum[4].

IV. INFORMATION FLOW IN AERIAL CONTOURING

Figure 2(P3) includes those parts of the aerial mapping process of interest here. The original radiance distributions, $s_1(x,y)$ and $s_2(x,y)$ are the radiances of the ground (suitably integrated over wavelength) as observed at stations #1 and #2, respectively. The relationship to the independent variables is: $s_1(x,y)$ is the radiance observed at station #1, in the direction of the point (x,y,o) which is in the datum plane, and similarly for $s_2(x,y)$. The variation with y is suppressed because of prior interior orientation and scanning as discussed above; one need consider only $s_1(x)$ and $s_2(x)$. If the terrain is diffusely radiating, i.e., radiance is independent of direction, if the pictures are taken sufficiently close together in time so that the shadows have not moved and if the terrain is level over the sub-image investigated, then $s_1(x)$ will differ from $s_2(x)$ only by a distance shift, or parallax. That is $s_1(x) = s_2(x + x_p)$ Usually, although the terrain is not diffuse, variations in radiance with angle have negligible effects. At the present and projected near future state of the art, when these variations are important, the skill and judgement of a human compiler are required. Thus, it is assumed that the terrain is diffuse, the shadows have not moved, and that a plotter with the capability to project the terrain, on the datum plane for measurement is in use.

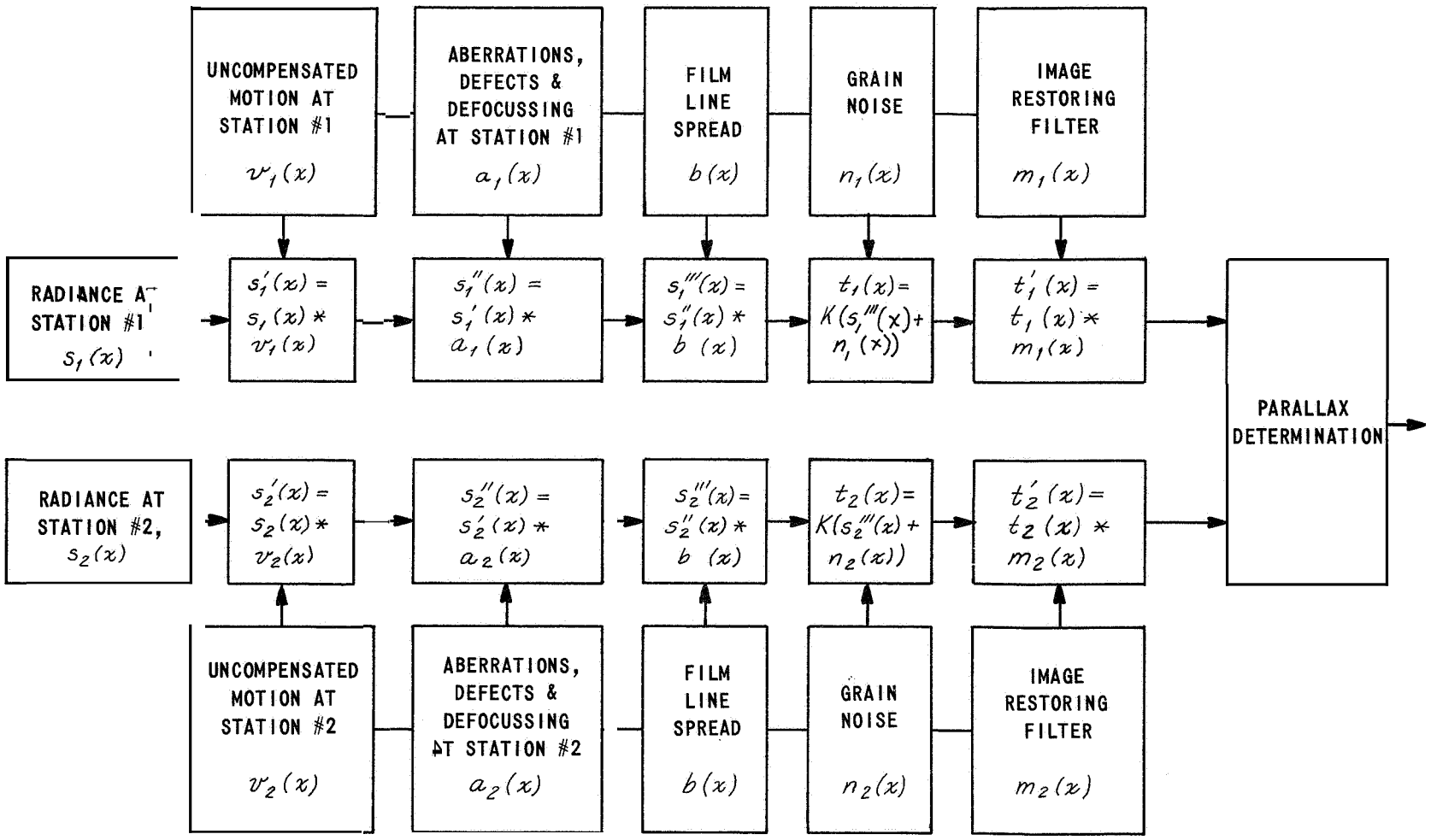


Figure 2 (P3) INFORMATION FLOW IN AERIAL CONTOURING, (PARTIAL)

The uncompensated image motion introduces a line spread. The radiance presented to the lens may validly be represented as a convolution, (*), of the desired radiance and the line spread. This motion may not be the same over the format, e. g., motion in yaw.

The aberrations, defects and defocusing of the lens cause an additional line spread, which is convolved with the radiance above. It is not valid to assign separate line spreads to these deficiencies; they must be considered as a whole. Usually these deficiencies will vary over the format, because coma, astigmatism, curvature of field, and lateral color are variations over the format. In this and the preceding paragraph, the line spread considered is the one applicable to that region of the format under consideration. In general, the aberration and defocusing contributions to line spread are dependent on wavelength, both because the diffraction theory of aberrations claims that the line spreads of monochromatic aberrations are functions of wavelength and because of secondary chromatic aberration. Assuming, however, that the geometrical theory of aberrations holds, i. e., large monochromatic aberrations, and that secondary and lateral chromatic aberrations are negligible, then a line spread may be assigned to the lens, independent of wavelength. If it exists, it is valid to convolve this line spread with the scene radiance function.

There is also a line spread associated with the film. It depends on the exposure and it may also depend on color. However, it is in those regions of the image where contrast is low that errors are most likely to be large, and here the exposure will be limited to a narrow range. Assuming that color variations are unimportant, a line spread may be assigned to the film. It also is to be convolved as those previously.

All of the assumptions of this section, except those concerned with diffuseness and shadows, are the standard assumptions of image detail processing.

Because the photographic process uses finite size silver halide crystals, it will be found that there are fluctuations in supposed uniform areas. For photogrammetric purposes, these fluctuations may be approxi-

mated by white, gaussian noise. This is called grain noise. It also depends on exposure. However, where contrast is low, it may be assumed constant, as in the previous sections.

Note that no further degradations are shown in the flow diagram. Usually more degradations are introduced in the printing of the diapositives. However, if image detail is so critical that restoration is being considered, a good first step would be to eliminate serious degradation in the laboratory, where exposure time is unlimited thus permitting the use of fine grain high resolution film and well corrected lenses at optimum aperture settings.

It is also to be noted that the transmittance of the diapositive is shown as being proportional to the radiance of the scene. This is strictly true only if the developing is done to make it so, i. e., the product of the γ's must be equal to 1. If this is not done, then a low contrast approximation is that the transmittance is proportional to radiance plus a constant which may be dropped. Justification for low contrast approximations was discussed previously.

V. RESTORATION FOR IDENTICALLY DEGRADED IMAGES

In this section, it is supposed that the degradations of both images are the same, and that the effects of data link are absent or negligible. Recalling that the target radiances were assumed the same in each image, it is seen that there are **two** functions:

$$\begin{aligned} t_1(x) &= K\left[s(x) * d(x) + n_1(x) \right] \\ t_2(x) &= K\left[s(x+x_p) * d(x) + n_2(x) \right] \end{aligned} \qquad \text{(P3-6)}$$

which are to be compared for parallax. In the above equations, the subscript on s has been dropped, and all degradations have been combined into $d(x)$ by interchanging the order of integration in the convolution integrals.

It may be argued that the real trouble arising from the noise is that it varies from image to image, and that if it were the same it could be compensated. From this point of view, $t_1(x)$ may be considered as perfect (with respect to noise) and $t_2(x)$ may be considered as corrupted by an additive white noise sample $K[n_2(x) - n_1(x)]$. However, from Section III above, if one wants a maximum likelihood estimate of the parallax of a corrupted function, then he convolves it with the time reversal of, **or** correlates it with, a perfect sample. Thus, correlation without filtering yields the maximum likelihood estimate. Any attempt to restore detail in the image would destroy the whiteness of the noise. Inspection shows that a colored noise filter formulation of parallax determination (required if image restoration were attempted) would undo the restoration and restore the degradations.

In the above discussion, the integrals were taken from minus infinity to plus infinity because the area over which the correlation is being performed is large compared to reciprocals of spatial frequencies of interest. However, the area is still small compared to the size of the diapositive. Because the correlation is computed over a limited area, no significant loss will result if, somewhere in the system, filters are inserted to remove modulation at frequencies well below the reciprocal of the dimensions of the correlation spot. Such a filter could remove lens falloff and vignetting, a source of error not usually considered in analyses such as this. Also, such "dodging" would reduce problems related to light level and dynamic range in diapositive preparation and use.

VI. RESTORATION FOR TRANSPARENCIES WITH DIFFERENT DEGRADATIONS

Assuming again that a data link is not of concern, a similar attack may be used if the degradations are different. Taking t_1 as perfect, t_2 may be expressed as

$$t_2(x+x_\rho) = K\left[\left(\frac{t_1(x)}{K} - n_1(x)\right) * d_1^{-1}(x) * d_2(x) + n_2(x)\right] \qquad \text{(P3-7)}$$

or

$$t_2(x+x_\rho) = t_1(x) * d_1^{-1}(x) * d_2(x) + K\left[n_2(x) - n_1(x) * d_1(x)^{-1} * d_2(x)\right] \qquad \text{(P3-8)}$$

where $d^{-1}(x)$ represents the inverse filtering operation, assumed to exist where needed. Thus, using the general colored noise formulation, the matched filter for t_2 is

$$M_2(\nu) = \frac{T_1^*(\nu)\, D^*(\nu)}{D_1^*(\nu)\, K^2\left[|N_2(\nu)|^2 + |N_1(\nu)|^2 \left|\frac{D_2(\nu)}{D_1(\nu)}\right|^2\right]} \qquad \text{(P3-9)}$$

where it is assumed that the grain is not correlated between diapositives. By using the fact that the noise is white, and removing constant factors one may arrive at

$$M_2(\nu) = \frac{D_1(\nu)\, T_1^*(\nu)\, D_2(\nu)^*}{|D_1(\nu)|^2 + |D_2(\nu)|^2} \qquad \text{(P3-10)}$$

or

$$M_2(\nu) = \left[D(\nu)\right] \cdot \left[T_1^*(\nu)\, D_2^*(\nu)\right] \qquad \frac{1}{|D_1(\nu)|^2 + |D_2(\nu)|^2} \qquad \text{(P3-11)}$$

A startling method of interpreting the above result is to suppose that each "restoring" filter gives to its transparency the degradations of the other transparency. Then the $T_1(\nu)^*\ D_2^*(\nu)$ term in the numerator represents correlating with the other (filtered) transparency, and the $D_1(\nu)$ term represents the degradations to the second transparency by its "restoring" filter. The denominator represents filtering to boost the gain at degraded frequencies and may be inserted wherever convenient. It may

be noted that if the numerator of this scheme is implemented with aperture filters, and the denominator by the use of coherent optical filtering, then no negative or imaginary parts of filters are required. The statements of the previous section regarding dodging apply to this one also.

VII. EXTENSION TO SYSTEMS WHERE A DATA LINK MUST BE CONSIDERED

Figure 3(P3) shows the information flow for one of the images in a system using a data link; the information flow for the other is similar. The degradations have been combined as discussed previously. The new things to be considered are that the image may be sampled and that the image will be further corrupted by noise while being sent through the data link. These complications require additional filters in the system.

Because the images are two-dimensional and because the data link has available only one independent variable, time, it is necessary to sample the image in at least one of the dimensions. Sampling the image in both dimensions would be advantageous in some cases because it opens the possibility of using special modulation schemes to reduce required transmitter average power. This discussion is limited to systems in which the image transmittance is not quantized; i. e., position may be discrete but the value corresponding to scene radiance is continuous. According to the diffraction theory of image formation, (applied to the camera), the image will contain no modulation at spatial frequencies above a definite limit. Thus, according to the sampling theorem, it may be reconstructed perfectly from equally and sufficiently closely spaced samples. However, the grain noise spectrum extends beyond the limit and thus a low pass filter is required prior to sampling lest the grain noise at higher frequencies be folded into the transmitted samples.

Although shown schematically as a single filter, this filter may be composite and complicated. In the case of photographic film being sampled by a flying spot or mechanical scanner, the filter is composed of the intensity distribution of the spot to be imaged, aberrations, defects, and defocusing at various places in the optical system and depending on the type of scanner, the sensitivity of the detector. Also, it may be desirable to use specially

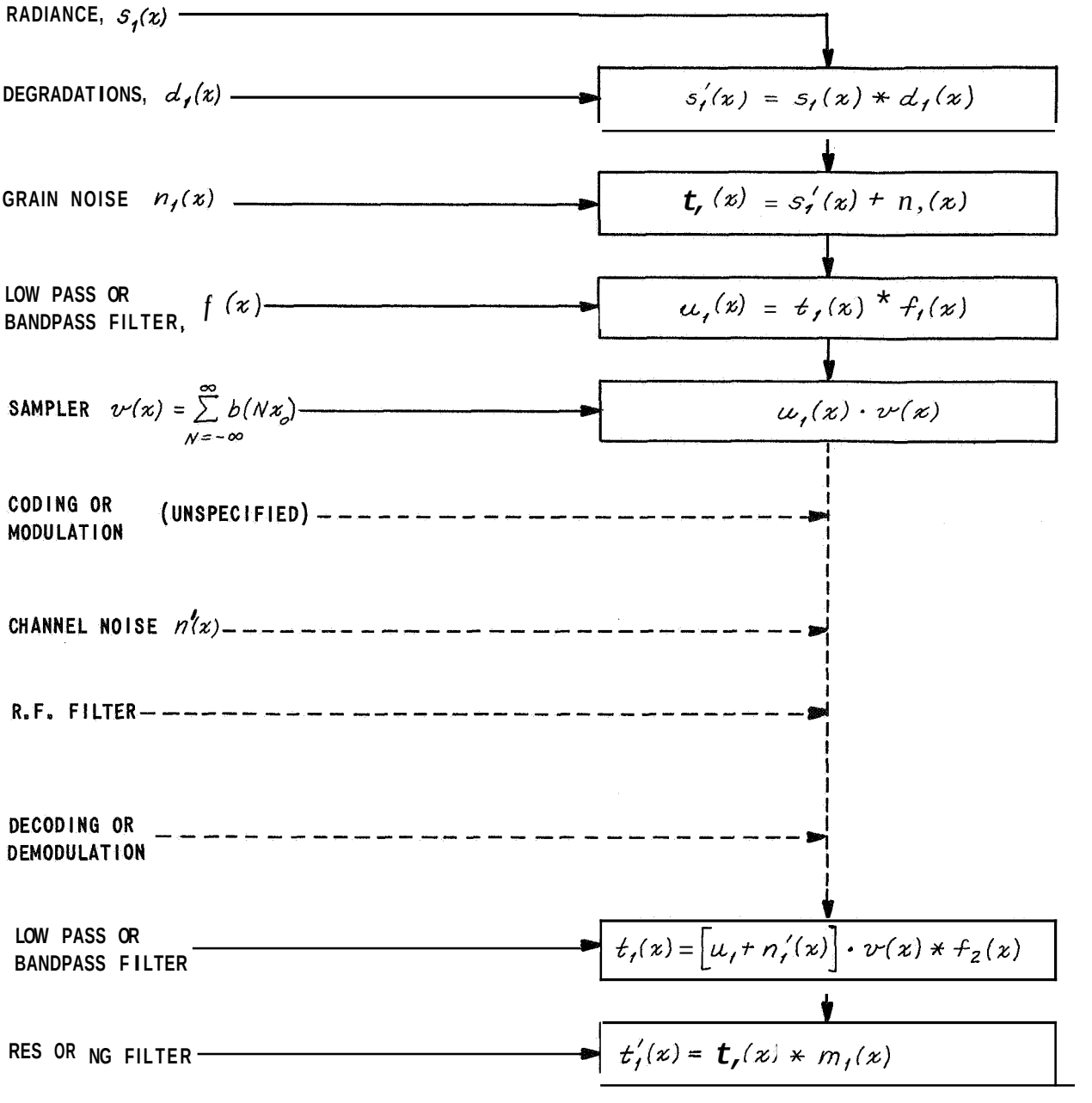


Figure 3 (P3) INFORMATION FLOW INCLUDING DATA LINK

shaped apertures over some of the lenses to create specific transfer functions, and to use numerical type filtering on the sampled data, such as could be provided by a tape loop with multiple read heads,

An interesting problem for further study is to find the optimum transfer function of this composite filter and to discover how closely it may be approximated at particular costs in weight and complexity. This is a valid question because, assuming optimum filtering at the compiler and reasonable approximations to signal spectrum, expressions may be derived giving mean square parallax error when the transfer functions of the degradations and composite transmitter filter and both grain and data link noise levels are known. These expressions could be evaluated for several pairs of values on a data link possibility curve; i. e., a curve of signal-to-noise ratio as a function of bandwidth for fixed weight and power consumption, or a convex linear combination of the latter two.

As discussed in Section V, there is no need to transmit extremely low frequencies. Also, there is no objection to a sharp cut-off because ringing caused by it is expected to be acceptable in an automatic plotter. Obviously, there is no need to transmit beyond the limit of the image, but the location of the high frequency cut-off and the shape of the transfer function up to it are not determined.

Summarizing the previous few paragraphs, the image is low pass filtered and perhaps sampled, depending on the type of system and direction under consideration. The next steps are modulation or coding, the unavoidable addition of noise, filtering at radio frequency, demodulation or detection, and further filtering. It is assumed that all of the steps of the previous sentence are equivalent to the addition of noise $n'(x)$ plus further linear filtering by a combined filter $f_2(x)$. This filter includes the effects of scanning spot size and of the line spread of the printing lens. Finally, the image may be operated on by a linear restoring filter before being correlated with the other image of the pair to determine parallax.

The most important task of the restoring filter, if it has not been performed by $f_2(x)$, is to remove the line-scan or halftone structure. If this is not done, the plotter will be sidetracked by this spurious detail rather than find the actual parallax. This rnay be seen from either a theoretical or practical standpoint. Practically, it is obvious that the line or halftone structure will result in many spurious correlation peaks. Theoretically, unless both images happen to be samples with the sampling points located exactly the same with respect to the image (impossible over the whole image for terrain with slopes), failure to remove the spurious structure does not agree with the idea of convolving with a space reversal. According to the sampling theorem, the effects of sampling on the signal (but not on the noise) may be completely removed by using a suitable sharp cut-off low pass filter at the receiver. Ringing in the reconstructed image can result from data link operations only if the transmitter filter has a sharp cut-off below the time frequency corresponding (through scanning speed) to the spatial frequency limit of the camera lens. Furthermore, as noted previously, ringing is not expected to be a hindrance to automated compilation.

The MTF of the restoring filter may be regarded as the product of the MTF of the above halftone or scan line stripping low pass filter, and the MTF of a filter with suitable variations in the low passband, which is to be derived below.

Assuming that low-pass filtering has removed the sampling effects, the diapositive transmittances are

$$t_1(x) = \left\{ \left[s(x) * d_1(x) + n_1(x) \right] * f_1(x) + n_1'(x) \right\} * f_2(x) \qquad \text{(P3-12)}$$

and

$$t_2(x + x_\rho) = \left\{ \left[s(x) * d_2(x) + n_2(x) \right] * f_1(x) + n_2'(x) \right\} * f_2(x) \qquad \text{(P3-13)}$$

By expressing t_2 in terms of t_1,

$$t_2(x+x_p) = t_1(x) * d_1^{-1}(x) * d_2(x) + \left\{ \left[n_2'(x) - n_1'(x) * d_1^{-1}(x) * d_2(x) \right] + \left[n_2(x) - n_1(x) * d_1^{-1}(x) * d_2(x) \right] * f_1(x) \right\} * f_2(x) \quad \text{(P3-14)}$$

one may arrive at the matched filter for the second transparency

$$M_2(\nu) = \frac{T_1^*(\nu)\, D_2^*(\nu)\, D_1(\nu)}{\left\{ \left[|N_2'(\nu)|^2 + |N_2(\nu)|^2 |F_1(\nu)|^2 \right] |D_1(\nu)|^2 + \left[|N_1'(\nu)|^2 + |N_1(\nu)|^2 |F_1(\nu)|^2 \right] |D_2(\nu)|^2 \right\} |F_2(\nu)|^2} \quad \text{(P3-15)}$$

Using the fact that the quadratic spectrum of the noise will be the same in both diapositives, this may be interpreted more easily as the product

$$M_2(\nu) = \left[\frac{T_1^*(\nu)\, D_2^*(\nu)\, D_1(\nu)}{|D_1(\nu)|^2 + |D_2(\nu)|^2} \right] \cdot \left[\frac{1}{|F_2(\nu)|^2} \right] \cdot \left[\frac{1}{|N'(\nu)|^2 + |N(\nu)|^2 |F_1(\nu)|^2} \right] \quad \text{(P3-16)}$$

The first term of this product appeared in the previous section and was discussed there. The second term requires the correction of, 'or inverse operation to, the receiver filter and the third term requires a partial inverse operation to the transmitter filter, depending on the ratio of grain noise to data link noise. It is to be emphasized that this filter function is only valid over the medium range of frequencies; there is no need to pass very low or high frequencies as discussed previously.

VIII. CONCLUSIONS

On the basis of the above study the following conclusions can be drawn when the data link need not be considered.

1) If the two images are differently degraded each by a single mechanism, so badly that the degradation may be taken as uniform over the format, then processing can help, although it has not been determined how much.

2) If the images have the same degradations, then processing is of no value.

3) It will be necessary to study the variation of transfer functions with format position and wavelength to determine the feasibility of improving results with "good" images, and

4) There will be no loss of accuracy from properly performed dodging.

In situations where the data link must be considered, all of the above conclusions apply as well as the following:

a) It is important to strip off the scan structure or halftone pattern,

b) A theoretical analysis of the transmitter filter function is indicated, and

c) The effects of quantized transmission on mapping accuracy should be studied.

IX. REFERENCES

1. Swerling, P., "Parameter Estimation Accuracy Formulas", IEEE Transactions on Information Theory, Vol. IT-10, No. 4, p. 306.

2. Helstrom, C. W., Statistical Theory of Signal Detection, Pergamon Press, New York, 1960.

3. Esten, R. D., et al, "Automation of Stereocompilation", Manual of Photogrammetry, American Society of Photogrammetry, Falls Church, Virginia, 1966, pp. 759-802.

4. Turin, G. L., "An Introduction to Matched Filters", IRE Transactions on Information Theory, Vol. IT-6, No. 3, June 1960, p. 317.

APPENDIX P-4

IMAGE RESTORATION EXAMPLES

By P. G. Roetling, R. C. Haas and D. A. Richards

This appendix presents some examples of detail restoration in motion blurred space imagery to demonstrate the potential value of image processing to the space effort. Photographic examples of enhanced imagery generated on an internal research program* of the Cornel1 Aeronautical Laboratory are presented with appropriate explanations of the techniques used to obtain them. It is cautioned that these represent first attempts at applying these methods to imagery degraded in situ and are not the optimized results that might be achieved.

I. INTRODUCTION

Due to a malfunction of the high resolution camera of the Lunar Orbiter I mission, imagery obtained was degraded by image motion blur. A sample of this imagery was obtained from NASA to experiment with restoration techniques. Two methods were attempted: coherent optical spatial filtering and incoherent differentiation. The following sections briefly describe these methods and give the results of the attempts to implement them for the degraded imagery mentioned.

II. METHODS

In previous work at Cornell Aeronautical Laboratory, a spatial filter had been fabricated for the removal of image motion. It approximates an inverse filter without attempting to whiten the image spectrum but rather it attempts to simulate the output of a well compensated incoherent optical

* Roetling, Paul G. and Roger C. Haas, "Final Report of the Study of the Feasibility of Enhancing Lunar Orbiter Images which have been Blurred by Image Motion", CAL Internal Research Authorization No. 85-176, 20 February 1967.

system. This filter was used to enhance (or deblur) the Lunar Orbiter I photograph obtained with the high resolution camera which had uncompensated image motion. It should be noted that the filter was the result of a single experimental attempt and is not an optimized filter.

The other method employed was a photographic implementation of the differentiation technique suggested by Harris-[*]. He utilized a computer for performing the necessary operations. In the photographic implementation a positive and a unit gamma negative transparency are first produced from the original blurred image. The two transparencies are then overlaid with a slight displacement in the direction of the image motion. A copy of this sandwich is then printed to a high gamma to restore the contrast of the original. This process is a differentiation in one direction of the density image of the original. Since the image motion was generated in exposure space rather than density space, this method only approximates the inversion of the motion process.

III. RESULTS

Figure 1(P4) shows a Lunar Orbiter I high resolution photograph which contains uncompensated motion. Figure 2(P4) shows the same photograph deblurred by the use of the spatial filter mentioned above (approximate inverse). This enhanced image looks generally sharper than the original, but it is also extremely noisy. One must be very careful not to interpret noise outputs as additional small craters on the lunar surface.

Figure 3(P4) is a similarly enhanced image but it was made from the negative of the transparency which was used as an original in the production of Figure 2(P4). The results are essentially the same.

---

[*] Harris, J. L., Sr., "Image Evaluation and Restoration", Journal of the Optical Society of America", Vol. 56, No. 5, May 1966, p. 569.

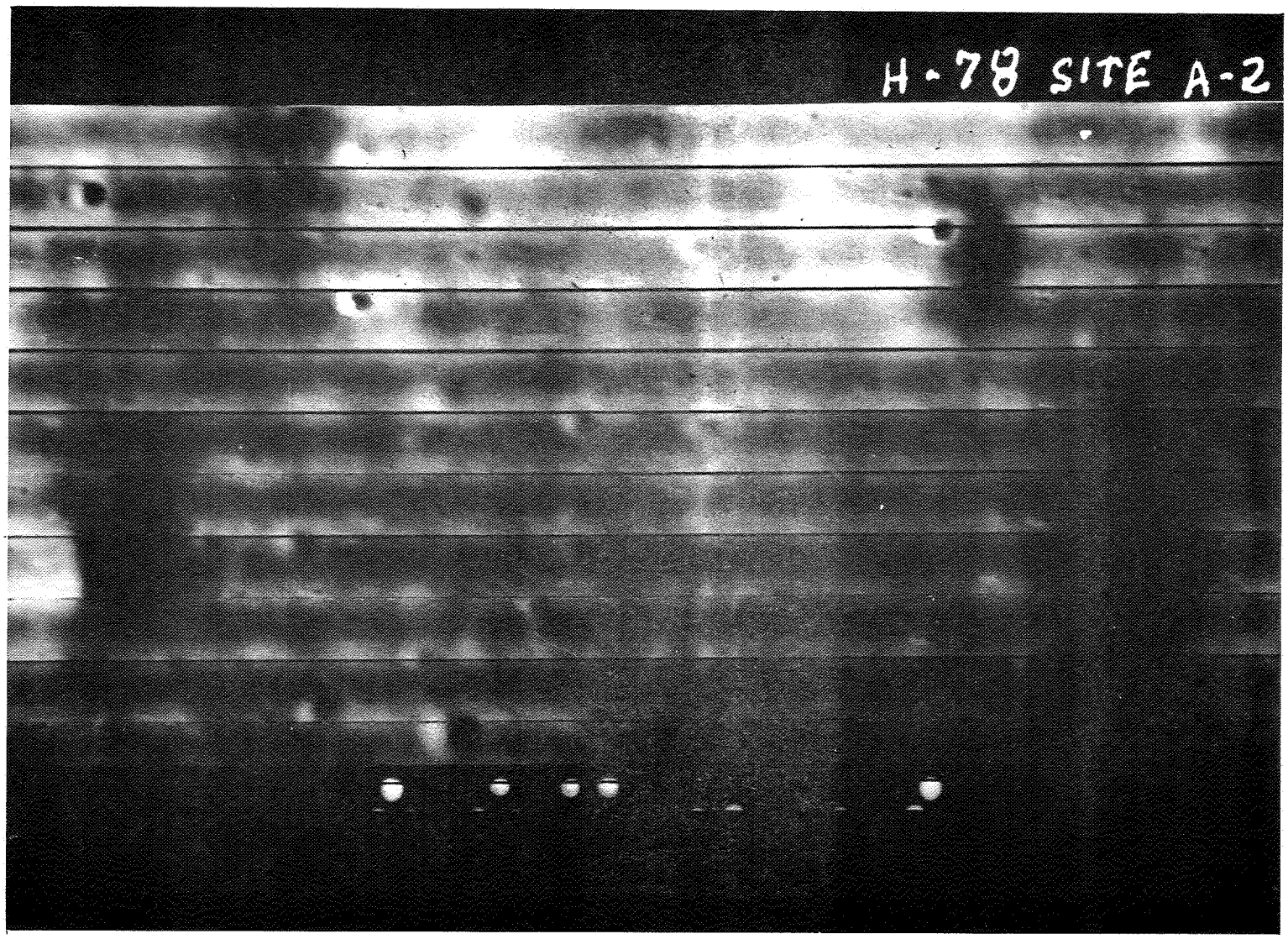


Figure 1 (P4) LUNAR ORBITER I HIGH RESOLUTION PHOTOGRAPH WHICH CONTAINS UNCOMPENSATED MOTION.

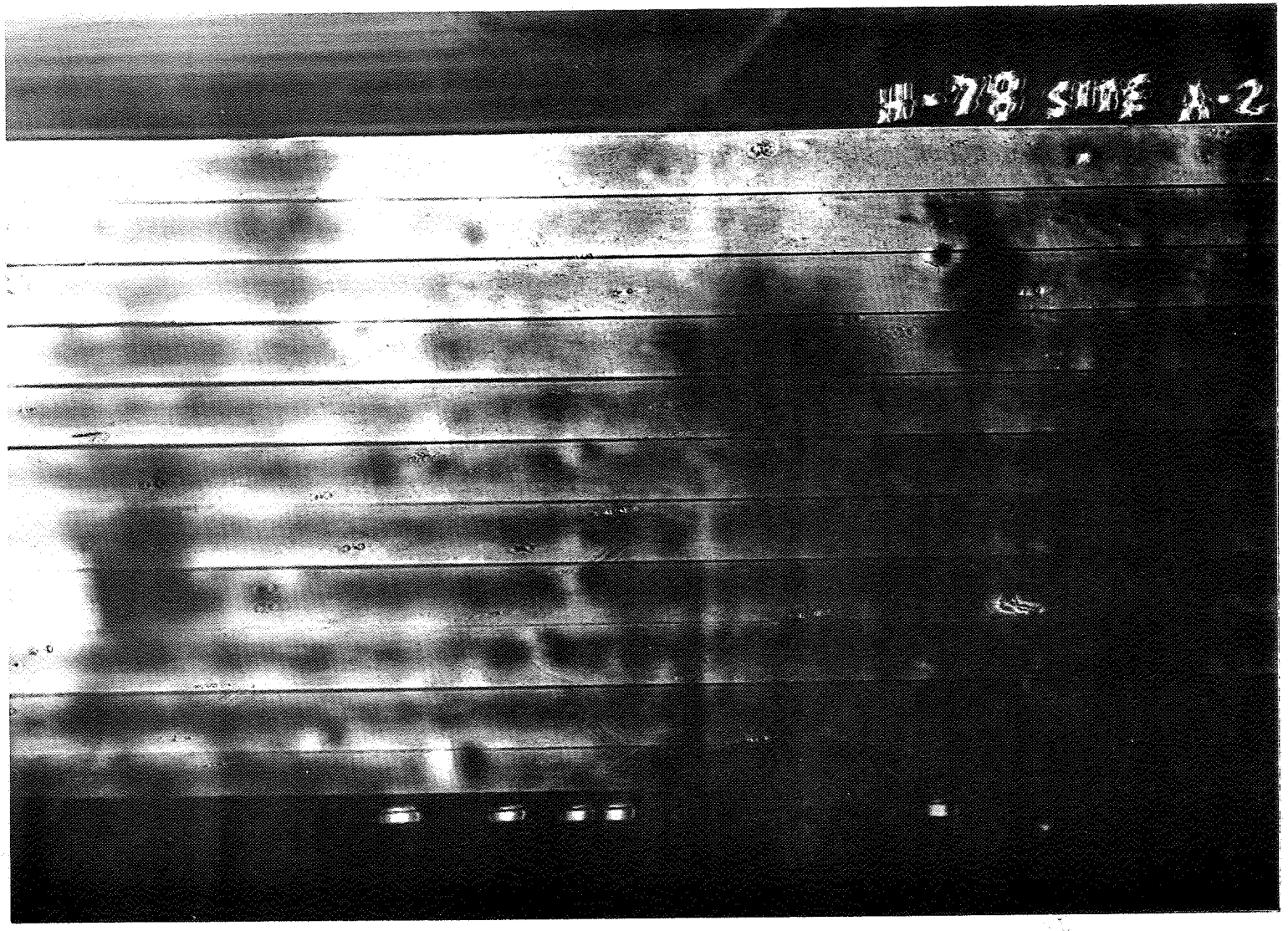


Figure 2 (P4) SAME PHOTOGRAPH DEBLURRED BY THE USE OF A SPATIAL FILTER (APPROXIMATE INVERSE).

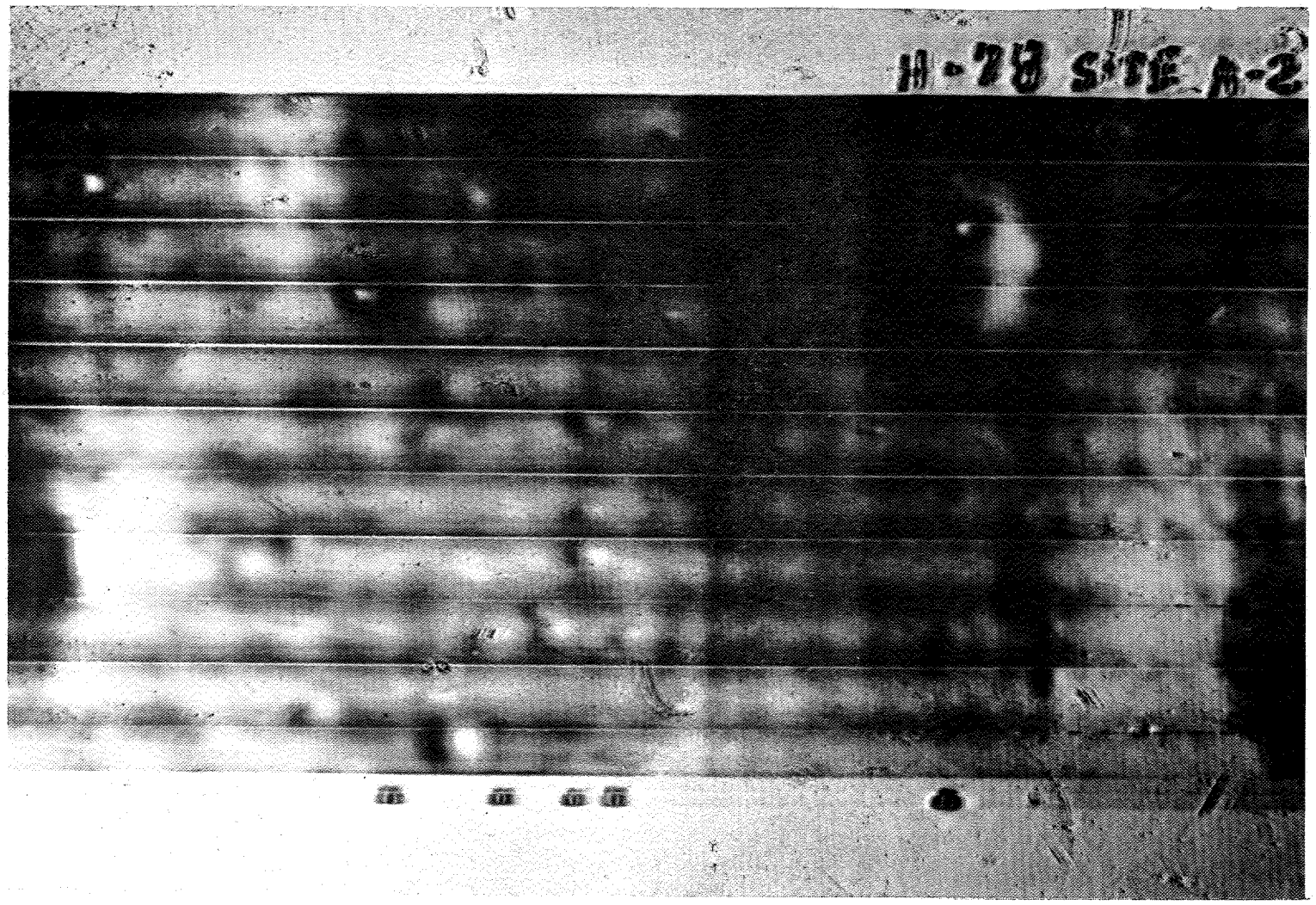


Figure 3 (P4) SIMILAR ENHANCEMENT MADE FROM NEGATIVE INSTEAD OF POSITIVE USED FOR FIGURE 2 (P4).

Figure 4(P4) shows the result of enhancing the photograph shown in Figure 1(P4) by the photographic implementation of the differentiation technique. It should be noted that this technique does not completely restore the image but rather creates a reasonably sharp positive image superimposed with a reasonably sharp negative image displaced by the amount of the image motion. With reasonable care, however, it is possible to interpret the two images separately and thus the result can be considered an enhancement of the original. The structured appearance of the general lunar surface as seen in this image is assumed to be due to the enhancement of noise rather than any real structure.

IV. REMARKS

These examples are included here only to illustrate some of the possibilities of image restoration. Their success with the Lunar Orbiter I imagery, although limited, demonstrates that these methods do have potential value and supports the arguments for more research and experimentation in this field. The development of a criterion for optimizing these methods appears to be the most critical problem.

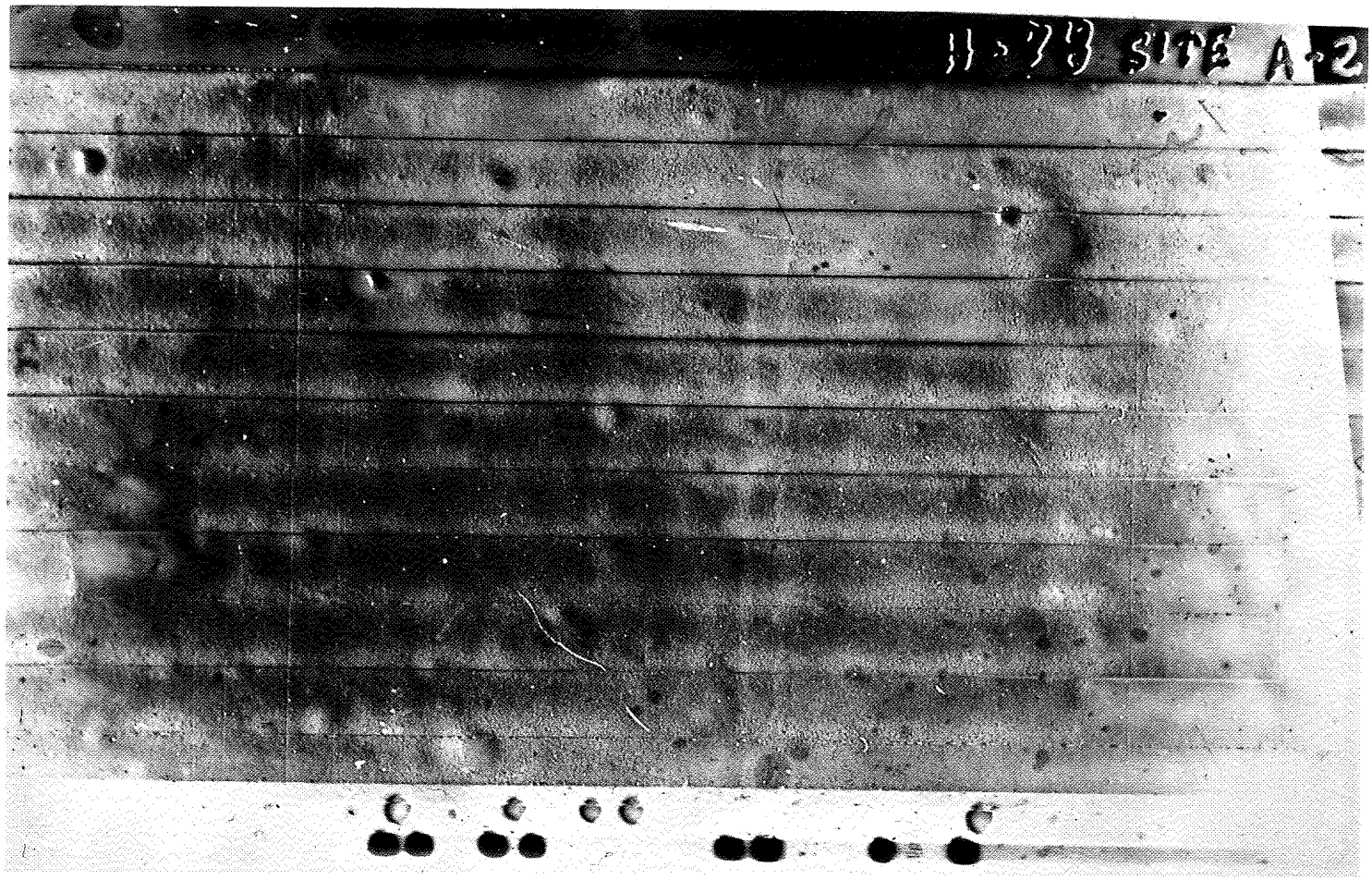


Figure 4 (P4) ENHANCEMENT BY PHOTOGRAPHIC IMPLEMENTATION OF THE DIFFERENTIATION TECHNIQUE SUGGESTED BY HARRIS (SEE TEXT).

APPENDIX P-5

EXPOSURE TIMES FOR STAR CAMERAS

By John D. Gallatin

In this appendix the length of time required to obtain an adequate exposure in a star camera and some of the resulting problems are discussed. Consideration is given to diffraction-limited systems and to a particular aberration-limited system. The analysis is limited to objects small compared to their distances from the camera.

I. INTRODUCTION

For planetary mapping from a space platform, a star camera has been recommended to determine the attitude of the platform (and cameras) for triangulation purposes, It is, therefore, appropriate to consider the exposure times necessary for such cameras. Two star camera systems are considered, namely, diffraction-limited and aberration-limited.

The motivation for considering diffraction-limited systems is: (1) it is desirable to use a sufficiently slow film to eliminate radiation shielding requirements, (2) it would then be advantageous to use a diffraction-limited lens to concentrate as much energy as possible onto a small area on the film. The motivation for considering aberration-limited systems is the exposure problem in the other type system.

In Section II the relationship between stellar magnitude and illumination at the entrance pupil of the camera is derived. This is used in Section III to find the required exposure time for diffraction-limited systems. In Section IV these results are applied to a particular system, and some of the resulting problems are discussed. A particular aberration-limited system is considered in Section V. The required exposure time for this system is found and the time and other characteristics are compared to the diffraction-limited system.

11. MAGNITUDE VS APERTURE ILLUMINANCE

The absolute magnitude of a star is a logarithmic measure of its intensity. That is

$$M = -2.5 \log_{10} \left(\frac{I}{I_o} \right) \qquad \text{(P5-1)}$$

where M = absolute visual magnitude

I = intensity of star in question

I_o = intensity of star with absolute magnitude of zero.

The constant I_o can be evaluated using the sun. From Reference 1 the solar illuminance at the earth's orbit, E_s, is 1.367×10^5 lumens/m^2, and from Reference 2, the absolute magnitude of the sun is t4. 8 and its distance from the earth, r, is 1.49×10^{11} meters. Since $I = E r^2$ the intensity of the sun is 3.04×10^{27} lumens/steradian or candles. Substitution into and rearrangement of the previous formula yields

I_o = 2. 52 x 10^{29} lumens per steradian for 0^{th} absolute magnitude star,

or $\log_{10} I_o$ = 29.402.

Reference 2 supplies the formula

$$M = m + 5 - 5 \log r \qquad \text{(P5-2)}$$

in which m is the apparent magnitude and for which r must be in parsecs. By using the fact that 1 parsec = 3.08×10^{16} meters, and other algebra, one can arrive at

$$0.4 M = 0.4 m - 2 \log r + 34.97 \qquad \text{(P5-3)}$$

for r in meters.

The result of combining the above formulas is

$$E_A = (2.67 \times 10^{-6}) \times 10^{-0.4m} \qquad \text{(P5-4)}$$

where E_A is the illuminance due to a star. The above is for a surface normal to the line of sight to the star just outside the earth's atmosphere and is approximately 20% greater than those generally quoted which apply at the earth's surface.

III. EXPOSURE TIME VS APERTURE ILLUMINANCE

For a diffraction-limited system the input luminous flux is the product of illuminance and aperture area. The detector illuminance may then be found by using the fact that all luminous flux incident on the aperture from a star will either be delivered to the star image or wasted. Causes of waste are obscuration of the beam by secondary mirrors and spiders, reflection from and absorption by lenses, windows, and filters, and absorption by mirror coatings. The image of a sufficiently distant star formed by a well-corrected optical system is a diffraction pattern. If the system is sufficiently well corrected, the illumination in the central part of the diffraction pattern will be essentially the same as that of a geometrically perfect system. According to physical optics, 83.8% of that portion of the luminous flux which is not wasted falls within the central spot of the diffraction pattern, or Airy disc, in a perfect system. Thus the luminous flux available at the detector may be formed by multiplying the input flux by 0.838 and by the efficiency, or the fraction of the flux not wasted.

The illuminance in the focal plane, averaged over the Airy disc, may be found by dividing the luminous flux by the area. The radius of the Airy disc (from physical optics) is

$$r = \frac{1.22 \lambda F}{D} \qquad \text{(P 5-5)}$$

where λ = wavelength
F = focal length
D = diameter of entrance pupil (assumed circular).

In order to express the radius this way rather than in terms of the size of the exit pupil and its distance from the focal plane, it is necessary to assume that the optical system satisfies the sine condition, or

(equivalently) is corrected for coma, which must be done for this type system. The statements and equations so far in this section may be combined into

$$E_F = \frac{0.838\,\varepsilon\,E_A D^4}{4(1.22\lambda F)^2} \qquad \text{(P5-6)}$$

where ε = efficiency of optical system
and E_F = illuminance on film.
Substitution of E_A from the previous section yields

$$E_F = \frac{0.838\,\varepsilon D^4 (2.67\times10^{-6})10^{-0.4m}}{4(1.22\lambda F)^2} \qquad \text{(P5-7)}$$

in which E_F will be given in lumens per square meter for λ, F, and D in the same units. For λ in microns, and F and D in inches

$$E_F = \frac{252\,\varepsilon D^4 \cdot 10^{-0.4m}}{\lambda^2 F^2} \qquad \text{(P5-8)}$$

As the exposure X is the product of the illuminance and the time, the time t rnay be found from: $X = E_F t$

$$t = X/E_F$$

$$t = \frac{X\lambda^2 F^2 \cdot 10^{0.4m}}{252\varepsilon D^4} \qquad \text{(P5-9)}$$

IV. NUMERICAL EXAMPLE OF DIFFRACTION-LIMITED SYSTEM

Assume a camera with 9" focal length and 3" aperture, or an f-number of 6, and an efficiency of 60%. Allowing an exposure of 0.32 lumen-sec/meter^2 (this corresponds to images visible with difficulty on SO-243 film under the most vigorous and extended development recommended), considering stars on the fourth magnitude and brighter, and taking $\lambda = 0.55\ \mu$ as a typical wavelength yields,

$$e = \frac{1}{10}\text{-sec.}$$

Of course, if this were a design of the camera and its use plan, a safety factor of 1-1/2 to 2 would be applied. Superficially, this appears to be a reasonable exposure time. But, further investigation reveals difficulties with I. M. C., stabilization and lens design.

Since the vehicle is orbiting with the topographical camera always pointed toward its primary, it is rotating with respect to sidereal space at the rate of one revolution per orbit. Taking the moon as an example, this rate is found to be about 9.6×10^{-4} radian/sec. Thus, if compensation were not used, and assuming no safety factor, the star image moves 96×10^{-6} radian. The angular radius of the star image is, by the use of previous formulas, about 9×10^{-6} radian. Thus, it appears that image motion compensation would be required for the moon. For primaries with shorter satellite periods, this requirement would be greater in terms of accuracy.

The stability required of the vehicle may be derived from the results of the previous paragraph.

As the exposure time is 1/10 sec and the image radius is 9×10^{-6} radian, the maximum random rotation rates in roll and pitch should be small compared to 90×10^{-6} radian/sec. This stability is thought to be beyond the state of the art.

There is also a serious problem in lens design. Inspection of star maps, Reference 2, shows that with care an orbit could be selected with nine each fourth magnitude or brighter stars per frame if the angular field were defined by format side equal to half of the focal length. Although a diffraction-limited f/6 lens is reasonable, a diffraction-limited lens to cover such a large field at f/6 is on the border of or beyond the state of the art.

Apparently, the only way to obtain adequate exposure is to use a much larger camera, with a slightly larger f-number lens, as the diameter appears to the fourth power but the focal length to only the second power. Reduction of the field of view to permit the use of a faster lens must be traded against the requirement to photograph dimmer stars with reasonable size cameras.

V. EXPOSURE TIME WITH A FASTER FILM AND ABERRATION-LIMITED LENS

The computation of exposure time with aberration-limited lenses and loss of contrast in the film cannot proceed in such a theoretical manner. In the absence of experimental tests, the image area must be arrived at by an educated guess based on resolution experiments.

Assume a system with 9" focal length and a 4-1/2" square format, using EK Plus-X Aerographic film, #5401. An empirical formula indicates that a resolution of 28 line pairs/millimeter can be obtained over the entire format at f/3, with a test object contrast of about 4 to 1. Thus, it appears that taking the image area as a circle with a diameter of 1/28 mm would insure capturing most of the light. The lens diameter is 76.2 mm resulting in a film illuminance

$$E_F = (76.2)^2(28)^2(0.60)(2.67 \times 10^{-4}) \times 10^{-1.6} = 0.182 \ \frac{lumens}{meter^2} \qquad \text{(P5-10)}$$

Again, assuming 60% efficiency and fourth magnitude stars, the exposure for this film, under extended development, is 0.0044 lumen-sec/meter^2. Hence, the required time is about 1/40 second.

It may also be found that the resulting angular image radius is 7.8×10^{-5} radian. By multiplying the angular rate for steady rotation derived in the previous section by the time of 1/40 sec, it may be seen that I.M.C. is not required. Also, since this same rate is reasonable for stabilization, it may be concluded that the stability problems are not insurmountable.

VI. CONCLUSIONS

Form the above considerations, it may be concluded that both systems have serious problems. The diffraction-limited system would be required to be very large, to use image motion compensation and to need a high degree of stabilization. The aberration-limited system by using a faster film becomes more susceptible to space-radiation fogging and consequently would require more shielding, a disadvantage in weight.

Further investigation should be performed: (a)to find the optimum field angle and implied required stellar magnitude; (b) to try other systems between the extremes of the examples. Although some improvement may result, it is believed that obtaining adequate exposure will result in a large and heavy camera or magazine.

VII. REFERENCES

1. Condon, E. V. and Odishaw, H., "Handbook of Physics", McGraw-Hill, 1958.

2. Baker, Robert A., "Astronomy", 8th Ed., D. Van Nostrand, 1964.

APPENDIX P-6

LASER RANGEFINDER CONSIDERATIONS

By J. D. Gallatin and D. A. Richards

Several complications arise in the analysis of the power requirements for a simple laser rangefinder system for use in a planetary mapping system. These requirements are derivable from system constraints and a maximum failure probability specified by the system designers. However, the analyses presented in the literature rely on an approximation which provides an underestimate with no true indication of how underdesigned the power requirement is. An approach has been devised whereby another approximation, which provides a higher estimate of required power and which, in conjunction with the other, brackets the exact answer. A better evaluation of the validity of power requirements calculations can now be made.

I. INTRODUCTION

During the First Phase of Project TECH TOP it was recommended that a laser rangefinder be used for obtaining scale for photography from a satellite[1]. Subsequently some preliminary analysis was performed concerning the power requirements for such a system[2]. However, the calculations made at that time utilized values of the parameters which gave a relatively large signal-to-noise ratio and resulted in the background being negligible compared to the signal. The question arises, what reduction in the power requirement is possible so that, although the noise is no longer negligible, the false alarm rate is tolerable for the intended mission? It is to this question that this appendix is directed.

11. THE PROBLEM

The system being considered consists of a pulsed laser transmitter and an optics-photomultiplier receiver. The receiver optics includes a narrow band filter centered on the laser wavelength and the photomultiplier is followed by thresholding circuitry for detection. Several complications

arise in attempting to derive the appropriate threshold setting and in turn the signal power requirement for a design with low false alarm* rate and small miss** rate.

In connection with a planetary mapping which required M separate photographs, one would require M distinct altitude determinations. The laser rangefinder mission then becomes the accurate measurement of M distances. Let us start with the probability of a mission success, $P_M(S)$ which is the probability of no false alarms occurring in M trials. This is determined by the system designer and can be represented as

$$P_M(S) = 1 - P_M(F) = \left[1 - P_\tau(FA)\right]^M \qquad (P6\text{-}1)$$

where $P_M(F)$ is the probability of a mission failure and $P_\tau(FA)$ is the probability of an individual determination false alarm in a range gate time τ . For small $P_\tau(FA)$

$$P_M(F) \simeq M P_\tau(FA) \qquad (P6\text{-}2)$$

The time our receiver is operable (T) can be divided into k cells of length equal to the return pulse length (Δt) which must be

$$\Delta t \geq \Delta t' \qquad (P6\text{-}3)$$

where $\Delta t'$ is the original laser pulse length. (The inequality is included because terrain will generally broaden the return pulse.) This allows

---

* A false alarm, the reporting of a signal when none is present, is caused by noise exceeding the threshold.

** A miss, the failure to detect a signal when one is present is caused by signal plus noise not exceeding the threshold.

description of the single determination false alarm in terms of an individual cell false alarm

$$P_T(FA) = 1 - \left[1 - P_c(FA)\right]^k \qquad \text{(P6-4)}$$

where $P_c(FA)$ is the probability of a false alarm in an individual cell.

We can consider the photomultiplier as a photoelectron counter which counts for a time τ . τ should be ideally chosen equal to the return pulse length a t . If no signal is present we expect to count m photoelectrons where

$$m = \dot{\bar{n}}_B \tau \qquad \text{(P6-5)}$$

and $\dot{\bar{n}}_B$ is the average rate of arrival of photoelectrons due to the background noise. When a signal is present m becomes

$$m = \dot{\bar{n}}_B \tau + \bar{n}_s \qquad \text{(P6-6)}$$

where $\bar{n}_s$ is the average number of photoelectrons in a time τ due to the signal. When the range cell coincides with the return pulse

$$\bar{n}_s = \dot{\bar{n}}_s \Delta t \propto \frac{P_T}{R^2} \qquad \text{(P6-7)}$$

where $\dot{\bar{n}}_s$ is the average rate **of** arrival of signal-generated photoelectrons. This is proportional to the original pulse power P_T and can be firmly connected to it by a combination of known system parameters and the range R .

In any given cell count, the probability of counting n photoelectrons when m is most probable is given by the Poisson expression

$$p(n) = \frac{m^n e^{-m}}{n!} \qquad \text{(P6-8)}$$

The total probability of exceeding a predetermined value n_T (threshold) is then given by

$$P(>n_T) = \sum_{n=n_T}^{\infty} p(n) = \sum_{n=n_T}^{\infty} \frac{m^n e^{-m}}{n!} \qquad \text{(P6-9)}$$

If the m in expression (P6-9) is replaced with (P6-5) we have the probability of a false alarm $P_c(FA)$. If we use (P6-6) in (P6-9) we have the probability of detection $P_c(D)$.

Thus from a $P_M(S)$, M, k and τ specified by the system designer, one can determine the necessary $P_c(FA)$ and subsequently n_T. Similarly, from a $P'_M(S)$ [defined as the probability of no missed signal detections in M trials

$$P'_M(S) = \left[P_T(D)\right]^M \qquad \text{(P6-10)}$$

where $P_T(D)$ is the probability of detection in a single determination and

$$P_T(D) = 1 - \left[1 - P_c(D)\right]^{k'} \qquad \text{(P6-11)}$$

analogous to equation (P6-4) where k' is now equal to one and

$$P_T(D) = P_c(D) \qquad \text{(P6-12)}$$

---

* $P'_M(S)$ can be less than $P_M(S)$ since misses are less important to the mapping problem than errors due to false alarms.

we can determine $P_c(D)$. Using n_T and $P_c(D)$, one arrives at an m in (P6-9). This is then used in conjunction with the other system parameters (aperture, transmittance, etc.) and a knowledge of the background noise to generate P_T, the laser power requirement.

III. THE DIFFICULTY

The major difficulty which is contained in the analysis of Section II above is the rather arbitrary selection of k independent cells. This implies that our detection system counts - decides - empties and then starts again, This is not the case in the real system which is continuously operable. What is needed is a sliding cell. To illustrate this point let us look at Figure 1(P6). There we see a plot of signal level against a time axis divided into cells of equal length τ. We see that in no cell does the average signal exceed the threshold level. However, if we define a new cell between the vertical arrows we have a possible signal present or a false alarm not considered in the theory presented above.

From this illustration we see that equation (P6-4) is only an approximation because it does not consider all the possibilities of a false alarm. Such an approximation leads to an underestimate of the power requirement.

IV. THE SOLUTION

To obtain a more rigorous solution to the problem of determining the power requirement for this laser application, many powerful methods of statistical decision theory were reviewed. None appeared applicable to eliminating the particular difficulty stated in Section III. It was at this point that another approximation was decided upon.

The main objection to equation (P6-4) is not that it is an approximation, but that it is not obvious how good an approximation it is. A procedure was arrived at by which another approximation would give a higher number of false alarms than actually occur and together with (P6-4) would bracket the rigorous solution. With such an analysis a determination of how good the original approximation was, can now be made.

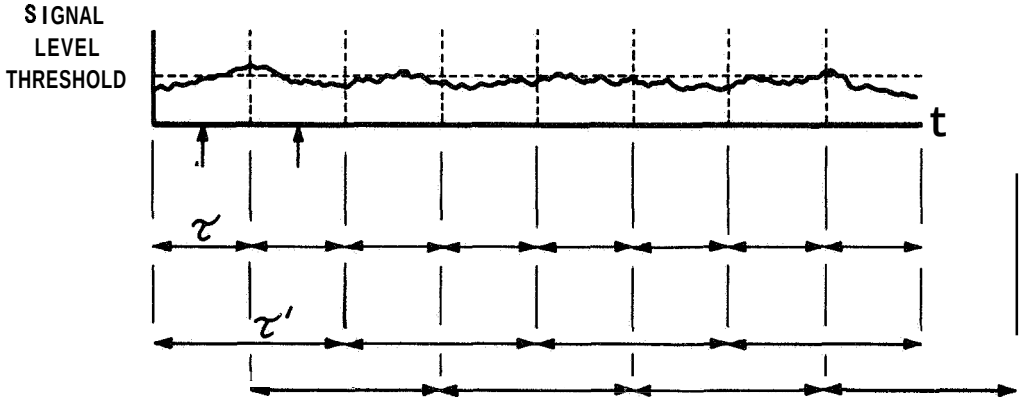


Figure 1 (P6) ILLUSTRATION OF RANGE CELL ANALYSIS

Referring once again to Figure 1(P6) we now take cells of length 2τ but such that they overlap from center to center. We then can consider k cells of length τ' where

$$\tau' = 2\tau \qquad \text{(P6-13)}$$

and use the same analysis presented above to arrive at an overestimate of the power requirement. By examining the figure carefully it can be seen that no false alarm escapes this analysis but also that it includes false alarms twice on some occasions.

V. EXAMPLE

Using a photoelectron count due to the background of 1.5 (calculated for a typical system applied to the moon) and the following values of the initial requirements

$$P_M(S) = 0.9999 \qquad M = 1000$$

$$\tau = 20 \text{ nanoseconds} \qquad k = 10^4$$

one obtains $n_T \simeq 10$. Combining this with a $P_M'(S) = 0.9$, results in a signal photoelectron requirement of 26.5 ($\overline{n}_s$) which in combination with other system parameters would provide the power requirement according to (P6-7).

Redoing the calculations with $\tau' = 40$ nanoseconds, one obtains a $\overline{n}_s' = 31$. Recalling that this is an overestimate and $\overline{n}_s = 26.5$ is an underestimate, we can see that a power requirement based on the first approximation would be lower than that actually needed by about 10%.

VI. REMARKS

It should be cautioned here that the results of calculations presented above are for a particular case and are included simply to numerically demonstrate the previous arguments. Each particular application of this analysis should be treated separately. The only general conclusion that can be drawn is that the two approximations did not differ by a great deal

and that minimal power requirements can be derived with greater confidence for this system. The reader is reminded that there are other systems of laser rangefinders (e. g., modulated cw) which have not been included in this investigation and would require their own analysis.

VII. REFERENCES

1. Campbell, C. E., "The Application of Laser Rangefinders to Ground Control for Photography from a Satellite", CAL Internal Memorandum dated 5 November 1965, as attached to Project TECH TOP (NAS12-19), Fourth Monthly Progress Report, October 1965.

2. Campbell, C. E., "Power Requirements for Laser Rangefinders", CAL Internal Memorandum dated 4 January 1966, as attached to Project TECH TOP (NAS12-19), Sixth Monthly Progress Report, December 1965.

APPENDIX S-1

DIFFERENTIAL HELIOROTATION

By Robert O. Breault

This appendix describes the phenomenon known as solar differential rotation. Three solar models are reviewed, supplemented by some derivations of pertinent equations. An empirical law which generates the solar rotations as a function of latitude is presented. Doppler shift and Doppler broadening techniques used in determining velocity and temperature are discussed, and the effect of the index of refraction of the solar atmosphere is also considered.

I. INTRODUCTION

The Sun does not rotate as a solid body, but rather, its period is a function of latitude. Observational data has determined that the angular velocity is largest at the solar equator and decreases toward the poles. This phenomenon is termed "differential rotation".

The problem presented by the differential rotation of the Sun may be stated most simply as follows. "The effect of viscosity within a mass of fluid such as the Sun acts in such a way as to reduce all internal relative motions and hence tends to make the mass rotate progressively more nearly as a solid. Since the most plausible hypothesis is that internal friction is present, there must also be present some systematic flow regime whose action, opposes the effect of viscosity, so as to maintain the observed differential rotation". [1]

A multitude of theories attempting to explain how the equatorial acceleration of the Sun originated or is maintained have been proposed. These include: Cowling's (1953) "fossil" rotation theory, which has the excess velocity left from the birth of the solar system; Babcock's (1961) proposal which depends upon the magnetic field to supply angular momentum to matter temporarily ejected from the sun; Randus's (1942) claim that rotating convective stars will necessarily have higher temperatures near

poles and states that an equatorial acceleration is required to counter the tendency for an unstable meridional cell to develop; Krogdahl (1944) attempted to show "that in viscous stars rotation will give rise to meridional currents, which in turn, cause a differential rotation". Schwarzschild (1947) derived a numerical solution for a slowly rotating star consisting of a convective core in solid-body rotation in a radiative envelope free of large scale convective currents, His solution is found to give an equatorial acceleration on the surface in qualitative agreement with the solar observation. Ward (1964) proposed that the equatorial acceleration of the solar photosphere is maintained against frictional and other retarding effects through the equatorward transportation of angular momentum from higher latitudes in each hemisphere. Ward's concept has been utilized and advanced by Starr and Gilman (1965) and feasibility estimates with respect to the original premises have been examined by MacDonald (1965).

MacDonald made a comparable study of Ward's methods in a test program of the terrestrial atmosphere. Using the large vortex motion of cyclones and anticyclones in 10° latitudinal belts analogous to the vortexial motion of sunspots, he concludes that "in a qualified overall manner, Ward's procedures for measuring the large-scale horizontal eddy components from sunspot displacements are given vindication".

R. Richardson and M. Schwarzschild have postulated[2] a very slow torsional oscillation superimposed on the differential rotation. The vibratory oscillation having a period of 22 years would thus explain the different rates at which sunspots migrate toward the equator in alternate 11-year cycles.

These are but a few of the past and contemporary models available for examination. In this work, those proposals which transmit the maximum information concerning fundamental solar physics have been subjectively extracted and, within limits, expanded upon. The hypotheses of Cowling (1953) and Babcock (1961) are briefly discussed. A relatively deeper penetration of Ward's hydrodynamical model experiments is presented because of the high degree of correlation between the theoretical results of the model and the empirical evidence available concerning the dynamical rotation of the sun.

The rotation of the sun, as defined according to its angular velocity, may be phenomenologically studied by a theoretical mechanism labeled the "law of isorotation". Under this principle, the sun is divided into cylindrical shells symmetric about the axis of rotation, with each rotating shell maintaining its own angular velocity and, if a general magnetic field is assumed, i. e., a polar field of the entire sun, wholly containing its own lines of force.

II. SOLAR MODEL I - GENERAL MAGNETIC FIELD

Ferraro (1937) has suggested the following explanation for the interaction of a general magnetic field and solar rotation. Since each shell contains its own lines of force, there exists an initial distortion in these lines of force when the angular velocity is not constant all along the line of force, i. e., each shell drags the magnetic lines of force along as it rotates. Since each shell rotates at a different angular velocity the lines become distorted. The distortion generates magnetic stresses, which, in turn react back on the motion analogous to the propagation of hydromagnetic waves. The effect is an irregular oscillation about the isorotational state. The period of oscillation then becomes comparable with the time of travel for hydromagnetic waves through the region of variable angular velocity. Conformable representation between irregular oscillations and hydromagnetic wave theory provides a simple and convenient model, but is inaccurate because of the unreliable correlation between the wavelength of the waves and the magnitude of the density variations.

The assumption of a surface magnetic field, such as found by Hale (1913), and an interior magnetic field increasing inward similar to Optical Depth formulation leads to two theoretical difficulties. Isorotation assumes a surface variation of the angular velocity whereas it has been indicated[3] that the angular velocity associated with each isorotational shell additionally varies with the height above the solar surface. Secondly, Walen[4] has pointed out, in the convective core the lines of force of a general field are likely to be tangled, and differences of angular velocity between separate lines of force are unlikely in the core. Thus isorotation outside the core would imply a uniform angular velocity at all points where lines of force from the core

reach the surface[5]. For the law of isorotation to have credibility, a force residing in the stable region of the sun, which is capable of appreciably altering the state of rotation since the time of formation of the solar system, must exist. The coercive constraint necessary to insure reliability of the isorotation law was thought to be that of a general magnetic field. Extensive measurements by Hale (1918) appeared to establish the existence of a polar field whose magnitude was about 25 gauss. Babcock (1948) reported various values for a general solar dipole field of 0 to 60 gauss. Recent comprehensive study has indicated that the field magnitude is much smaller than that reported by Babcock and that a general magnetic field is either non-existent or variable.

III. SOLAR MODEL II - MERIDIAN CIRCULATION

In the absence of a general magnetic field, a formulation designated as meridian circulation, may be used to describe the sun. The approach is that of a uniformly rotating star. Cylindrical polar coordinates r, θ, and z, with the axis of rotation as the polar axis and the center of the star as the pole (origin). Then, in uniform rotation with angular velocity $\omega = d\theta/dt$, the r and z equations of motion, as derived in Appendix S-1A, become:

$$\frac{\partial P}{\partial r} = (g_r + r\omega^2)\rho \qquad \text{(S1-1)}$$

where g_r is the component of gravity in the r direction, and

$$\frac{\partial P}{\partial z} = g_z \rho \qquad \text{(S1-2)}$$

where g_z is the component gravity of the z-direction. This fact becomes consistent with radiative equilibrium only if the rate of generation of energy per unit mass, ϵ, is represented by

$$\epsilon = \text{CONST.} \left(1 - \frac{P}{\pi G \rho}\right) \qquad \text{(S1-3)}$$

where P is the pressure, ρ is the density and G is the gravitational constant. (See Appendix S-1B).

The equations of motion satisfy the conditions of radiative equilibrium only on an average over a surface of constant gravitational potential, but the radiative condition is not satisfied at individual points of a level surface unless the energy generation rate equation is obeyed. Under these conditions, radiation produces heating near the pole and cooling near the equator. The hot material tends to rise, with cooler material sinking, resulting in circulating mass currents in meridian planes when the energy-generation equality does not hold. The speed of circulation has been found to be very small. When the viscosity of the material is considered, it is found that the viscosity tends to equalize the angular velocity throughout the sun. Similarly, the meridian circulation also tends to equalize the angular momentum per unit mass, but does not seriously modify the initial distribution of the angular velocity until the material has been transported a distance which is an appreciable fraction of the sun's radius. Under these conditions the circulation velocity has been found to be too small to be held responsible for the faster rotation of the sun at the equator than at high latitudes.

An alternative that would ideally conform to the observational data would exist if the meridian circulation velocity were much faster than calculated, i. e., greater than 4×10^9 cm/sec. Then the effects of rotation in the core would be of sufficient magnitude so that the convective transport of energy near the equatorial plane would be appreciably less than that near the axis. In addition, the temperature distribution outside the core would have to be correspondingly modified. Collective empirical knowledge on the physical state of the sun regards such a hypothesis as impossible. A present acceptable explanation is that the sun's equatorial acceleration is

either a relic of the past, i. e., a "fossil" rotation indicating its state when the solar system was first formed, or a consequence of a more recent capture, near the equator, of a large amount of material with a high angular momentum[4].

IV. SOLAR MODEL III - HYDRODYNAMICAL MODEL

All previous theories have assumed that the large scale fluid motions are symmetric about the rotational axis. Ward[6], in hydrodynamical model experiments has shown how rotation can significantly alter the fluid motions. In considering solar rotation and the magnitude and character of its latitudinal gradient, a hypothesis has been proposed holding that the general circulation of the solar atmosphere is not axially symmetric, but consists of wave motions, in which large components of the fluid velocities are along spherical surfaces. These waves (or eddies) transport angular momentum equatorward from high latitudes along these spherical surfaces by horizontal exchange processes.

Under steady-state conditions, the equations of motion (along a spherical surface in a coordinate system rotating with the Sun, and neglecting forces due to magnetic fields, etc.) show that there must be a balance between pressure forces and the centripetal and Coriolis accelerations. The ratio between the size of the acceleration terms in the equations of motion as compared with the Coriolis acceleration is known as the Rossby number. When the Rossby number is small, the pressure forces are balanced by the Coriolis force. When the Rossby number is large, the Coriolis term becomes negligible compared to the centripetal-acceleration term. This latter term must then balance the pressure force. The magnitude of the large-scale velocities in the Sun's atmosphere indicate the size of the centripetal effects, while the magnitude of the solar rotation itself is a measure of the Coriolis effect. The large-scale velocities can be estimated from either the characteristic sizes of the day-to-day proper motion of sunspots in the solar atmosphere, or by the magnitude of the latitudinal variation of the rotation.

In a previous study[7], the Greenwich photoheliographic data for the period 1935-1944 were examined for a correlation between the daily latitudinal and longitudinal proper motions of sunspot groups. The 10 years of spot data showed a very significant correlation of spot group proper motion, and in the proper sense. In a second study (Ward 1965), again the daily heliographic latitude and longitude of the Greenwich sunspot groups, for the years 1925-1954, were tabulated once each day. Here the rotation rate of 14.184° per day was assumed invariant with latitude. It was found that the position data, and hence the daily proper motions, are subject to relatively large and irregular fluctuations near the solar limb. Elimination of data obtained near the solar limb constituted 29% of that available.

In utilizing sunspots as tracers certain small scale discrepancies must be taken into account. Since the spots are circulations (or eddies), interactions between eddies of various sizes exist and presently are unknown. Additionally, the assumption that spots are imbedded in, and move with, the larger-scale flow field is not strictly valid. Moreover, non-random eddy positions with respect to the general circulation could lead to observational interpretations which are inaccurate in the determination of both latitudinal drifts and thus, the rotation rate. If one includes the probability that sunspots have a pronounced and systematically preferential radial tilt, and the combination of this tilt with the tilt of the solar axis and/or the brightness distribution of the spot as a function of height[2], then the production of spurious correlations, when studying the general circulation, yields only a first approximation.

Ward (1965) has concluded[6], "subject to the observational limitations outlined above, these data provide strong statistical evidence that the layer of the solar atmosphere in which spots are imbedded, extending presumably from below the visible surface into the low to mid-chromosphere, has a non-axially symmetric general circulation pattern of the Rossby type'.'. A Rossby regime is exemplified by fluid motions in which pressure forces are mainly balanced by Coriolis effects. The motions of the fluid are essentially horizontal, i. e., along spherical surfaces, and along these surfaces, the stream lines meander in a wavelike pattern, alternately poleward and

equatorward, but on the average progessively in one direction in longitude. The motions of the fluid through this wave pattern can be an order of magnitude faster than the longitudinal progression of the wave pattern itself. The motion of the wave pattern can be either eastward or westward, or stationary in longitude, with smaller oscillations in latitude. In all likelihood, each hemisphere of the Sun would have one or more such patterns with the possibility of coupling between hemispheres an open question. The number of waves, through 360" of longitude, their latitudinal extent, and their motions are also undetermined.

These observations strongly indicate that the so-called equatorial acceleration of the Sun has a different character from that usually visualized. It would now appear that there is an "equatorial acceleration" only if the data are averaged in time and longitude. Furthermore, this mean "equatorial acceleration" is maintained against an as yet unspecified dissipative force by the transport of angular momentum from regions in middle or high latitudes by selecting horizontal eddy processes. Since there is no way, at present, of estimating the magnitude of the flux of angular momentum required to maintain the observed equatorial acceleration, it is not certain whether this eddy flux is the primary (or the only) maintenance mechanism.

If one were to accept the Ward hypothesis, then the theories of Cowling and Babcock, previously mentioned, would become unfeasible. Cowling's conclusion of a "fossil" rotation would be destroyed since, if the source of momentum were suddenly shut off, with other processes continuing unabated, the layer of the Sun between the latitudes $\pm 27 - 1/2°$ would reach solid rotation in about 103 days, or four solar rotations. Likewise the calculations of Babcock show that the generation of the local magnetic fields around sunspots by his mechanism would exhaust all of the energy of the equatorial acceleration in about 1000 years.

Figure 1(S1) shows the sidereal rotation rate derived by Newton and Nunn[8] (1951) from recurrent sunspot group data. The circled dots in the figure show the rotation rate derived from the current data', which include not only the daily changes in the large, long-lived, recurrent groups,

but also the changes associated with smaller groups of shorter duration. Examination of these data reveal that the rotation rate derived from Newton and Nunn's recurrent group data is substantially and significantly lower than the rotation rate derived from all spot-group data. The small scattering among the individual cycles in Newton and Nunn's data show that the difference cannot be explained as a result of the period of time under consideration.

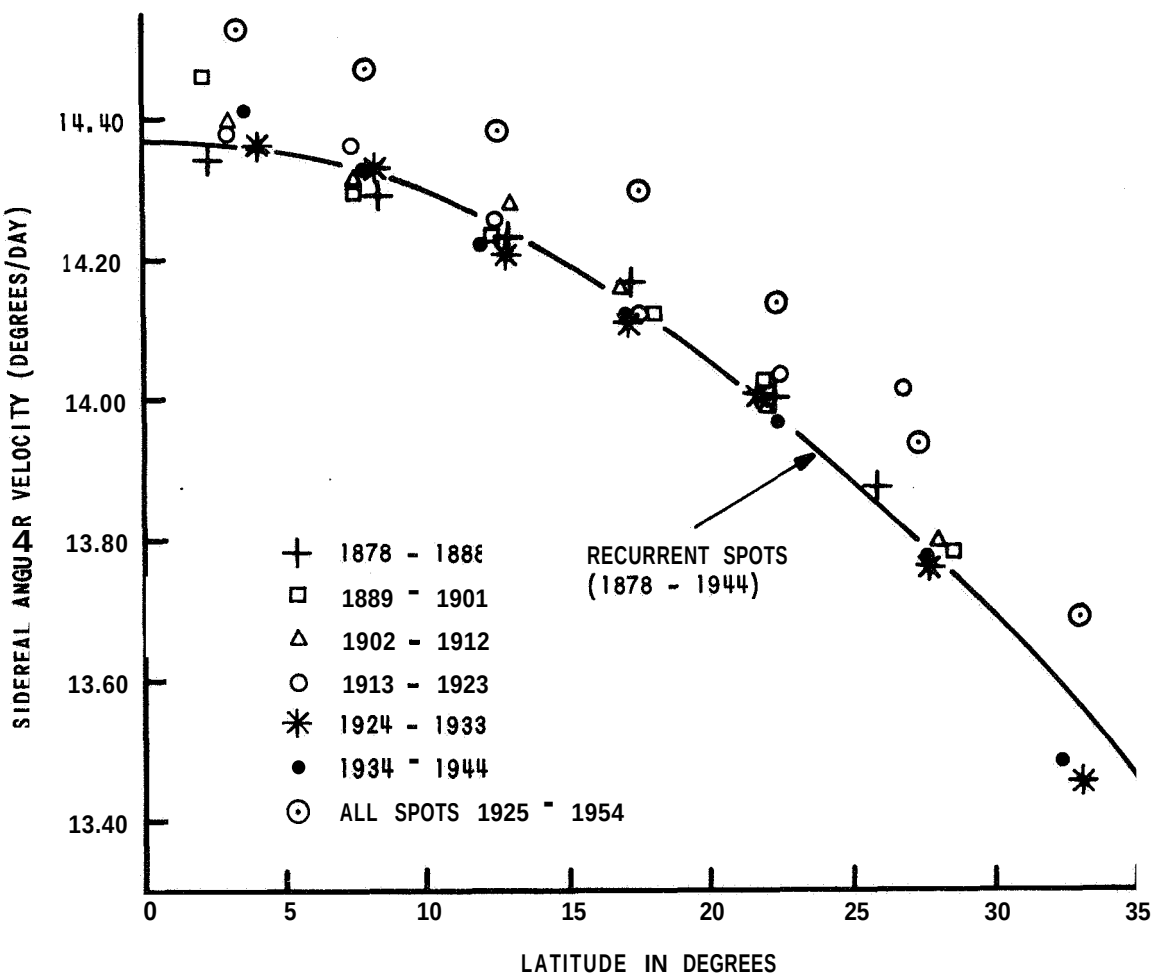


Figure 1 (S1) SOLAR ROTATION RATE AS DERIVED BY A NUMBER OF AUTHORS FROM SUNSPOT DATA (REF. 6)

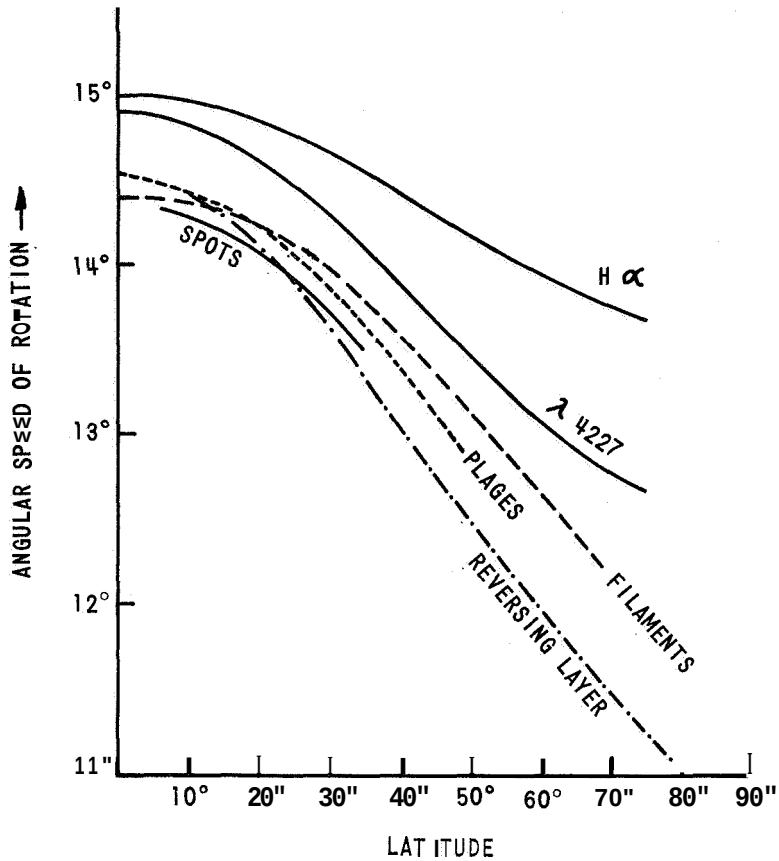


Figure 2 (S1) THE ROTATION OF THE SUN

MEASUREMENTS OF THE SOLAR ROTATION FROM THE POSITION OF SPOTS ON THE DISK WERE MADE MANY YEARS AGO BY SCHEINER, CARRINGTON, SPORER, FAYE, AND OTHERS. CHEVALIER AND MAUNDER OBSERVED THE ROTATION OF THE FACULAE, AND FOX MEASURED THAT OF THE FLOCCULI FROM THE SPECTROHELIOGRAMS OBTAINED AT YERKES. THE D'AZAINBUJAS HAVE DETERMINED THE SOLAR ROTATION FROM THE LONG-LIVED FILAMENTS. W.S. ADAMS, DUNER, AND OTHERS CARRIED OUT EXTENSIVE MEASURES OF THE SOLAR ROTATION FROM THE DOPPLER SHIFT IN THE FRAUNHOFER SPECTRUM AT THE LIMB OF THE SUN. (REFERENCE 9.)

The small differences may be ascribed to the variation in height or depth in the solar atmosphere through which large and small spots extend. Since the data includes many small spots, a tendency for these spots to move more rapidly would weight the results in the proper direction. Figure 2(S1) is a summary of results in determining the rotation rate from examination of spots, plages, filaments, reversing layer and Doppler measurements. The data indicates a tendency for faster rotation in the chromosphere and corona[9].

V. EMPIRICAL LAW OF ROTATION

Carrington (1863), from his visual observations of sunspots, invented a simple empirical formula which gives the angular velocity (ω) and hence the period of rotation (T) as the sum of a constant term and one that depends on the latitude. He found

$$\omega = 14^\circ.42 - 2^\circ.75 \sin^{7/4}\phi \qquad \text{(S1-4)}$$

where ϕ is the solar latitude, and ω is actually in degrees per day.

In 1951, Newton and Nunn summarized the results of six sunspot cycles, 1878 to 1944. They found no variation in the sun's rotational period, either with alternate 11 year cycles **or** within the cycle itself, and obtained **for** the law of rotation

$$\omega = 14^\circ.38 - 2^\circ.77 \sin^2\phi \qquad \text{(S1-5)}$$

in degrees per day.

The calculations for the period[*]. T are summarized in Table I(S1) below.

TABLE I(S1)

SOLAR ROTATION (DAYS) AS A FUNCTION OF SOLAR LATITUDE (DEGREES)

Latitude	0	5	10	15	20	27-1/2	30	35	40	45
Period	25.03	25.07	25.18	25.36	25.61	25.76	26.32	26.17	27.12	27.70

[*] The period is given in sidereal time, i. e., the time required for an object to move from a particular position among the stars back to the same longitude again, as seen from the sun.

In Figure 2(S1), the rotational velocity determined by means of radial velocity observations of Fraunhofer lines or emission lines show considerable scattering. At ϕ = 0", the ω -determinations by means of metal absorption lines in the photosphere range between 13". 2 and **14**". 8. The mean scatter is 0°. 35, hence 2. 5%, whereas the mean scatter derived from the observed motions of other various phenomena [Figure 2(S1)] is about 0. 4%, nearly one order smaller. At ϕ = 40" the scatter becomes twice that at ϕ = 0°. The explanation is: "the mean equatorial velocity is 2 km/sec; an error of 2. 5% corresponds to an error in the measured velocity of 0. 05 km/sec, hence with $\Delta\lambda = \pm 0.0008$Å. This is about the limit of accuracy of wavelength measurements". [17]

VI. DOPPLER EFFECTS

The empirical "law of rotation" discussed in the previous section generates a large, but finite, number of isomorphic solutions between the latitudinal coordinates and the angular velocity, but has experimental bounds in the form of an upper limit of 45" since sunspots rarely occur in higher latitudes. Spectroscopic verification and empirical datum expansion on the solar rotation, from an equatorial plane to the polar latitudes, are obtained utilizing the Doppler Principle. This principle provides two important physical parameters. The velocity of latitudinal rotation and the temperature distribution as a function of height above the photospheric surface are derived from the shifting and broadening of the spectral lines, respectively. Dark lines (Fraunhofer) appear in the spectrum of sunlight. These lines are caused by atoms of various gases located in the reversing layer (region of solar atmosphere where temperature gradient reverses sign).

Doppler Shift - An observer viewing a stationary absorption line will see the same frequency which is emitted by the source. If the source is moving toward or away from the observer, the frequency will be slightly changed by an amount that is proportional to the velocity. The frequency is increased - shifted toward the blue end of the spectrum - if the source is moving toward the observer and decreased - shifted toward the red end of the spectrum - if the source is moving away from the observer. The

applicable representation for light waves, depends only on the relative velocity of the source and the observer and is given as[10]

$$f = f' \frac{\left(1 - \frac{v}{c}\right)}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \text{(S1-6)}$$

where f is the observed frequency, in cycles per second, f' is the source frequency, v the relative velocity of the source, and c the speed of light.

Rewriting the equality in the coordinates of the observer's frame of reference

$$f' = f \frac{\sqrt{1 - \frac{v^2}{c^2}}}{\left(1 - \frac{v}{c}\right)} = f\left(1 + \frac{v}{c} + \frac{1}{2}\frac{v^2}{c^2} + \cdots\right) \qquad \text{(S1-7)}$$

This is the Doppler effect for the source and observers approaching each other with a relative velocity v along the line joining them. When comparison of the series expansion is made to the classical theory prediction, relativity differs only in the terms of second and higher orders. Additionally, relativity predicts a second-order Doppler shift even when the source is moving at right angles to the line of sight[11].

Dropping second and higher order terms and allowing $f - f' = \Delta f$ $\lambda = c/f$, results in the familiar classical representation

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c} \qquad \text{(S1-8)}$$

It should be noted that only the shift of spectral lines for velocities radially along the line of sight from the source to the observer will be detected. Those lines for which there is no velocity component in the direction of the observer will display a frequency of light equal to the actual frequency.

Because of rotation, the eastern equatorial limb of the sun is moving toward the observer with a speed of about two kilometers per second; likewise the western limb is receding at the same speed. The velocities produce a measurable shift of 1/150,000th part of the particular wavelength employed. The actual linear amount measurable is dictated by the dispersion of the spectrum and, additionally, is found to be a function of the wavelength of the Fraunhofer line used.

An important benefit derived by employing Fraunhofer lines is that those lines representing different elements at differing heights in the solar atmosphere provide information as to the velocity distribution as a function of height. The spectral lines of iron are found near the bottom of the chromosphere and those of hydrogen and ionized calcium extend upward. Hence, the sun's rotation may be studied not only from equator to pole, but also upward from the reversing layer to the upper levels of the chromosphere. Some conclusions drawn from Doppler shift observations are: in the reversing layer, the rotation period continues to increase up to the polar regions, where the period is about 30 days. At any given latitude, the rotation period decreases with the height above the reversing layer. At the equator, the rotational period drops from 25. 03 days to 23 days in the highest parts of the chromosphere. Thus, in the chromosphere there appears to be a westward shift as opposed to the eastward drift of the photosphere. At these higher levels, the rotational period of the higher latitudes becomes approximately equal to that at the equator. Coronagraphic data gathering shows the corona to have an approximate solid body rotational period of 25 to 26 days.

Line Broadening - The geometrical shape of a spectral line is influenced by the experimental instrumentation used to observe the source. In particular, the design of the slit of the spectroscope produces the rectangular shaped spectral lines. By reducing the slit width, higher resolution may be obtained, but no spectral line observed is perfectly sharp, no matter how great the resolving power of the spectrometer[**]. A line devoid of structure appears densest in the center and fades out symmetrically on the edges, in a gaussian form [Figure 3(S1)].

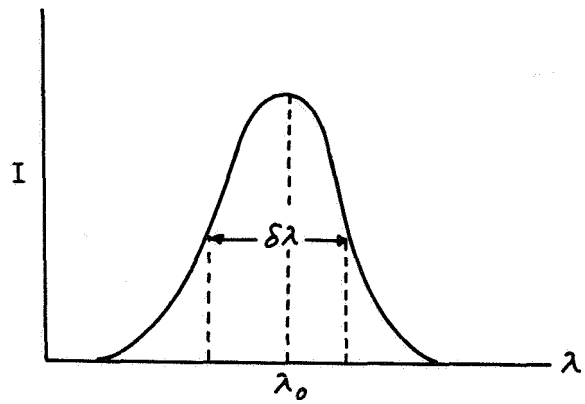


Figure 3 (S1) SPECTRAL LINE WIDTH

The width $\delta\lambda$ of a line is defined to be the distance between two points, one on each side of the center, at which the intensity, I, is half as great as it is in the center of the line. The primary causes of line broadening are due to the following mechanisms.

(a) Doppler Broadening - In a luminous gas the atoms are moving with a Maxwellian distribution of velocity, the average velocity increasing with temperature. A spectral line emitted by a gas thus comprises a range of frequencies symmetrically distributed about the frequency emitted by the atom when it is not in a state of thermal agitation. The distribution of intensity throughout the line is determined by Maxwell's distribution of velocities and is given by the equation[12],

$$I_f = I_{f_o} \exp\left[-\frac{Mc^2}{2kT\nu_o^2}(f-f_o^2)\right] \qquad \text{(S1-9)}$$

or

$$I \propto e^{-\beta\overline{\Delta\nu}^2}$$

An extension of (S1-9) gives the line width at half intensity, if the broadening is entirely due to Doppler effects

$$\delta\lambda = \frac{1.67}{c}\lambda_0\sqrt{\frac{2RT}{M}} \qquad \text{(S1-10)}$$

Substitution of the proper constants give

$$\delta\lambda = 0.72 \times 10^{-6}\lambda_0\sqrt{\frac{T}{M}}$$

in wavelength units, where M is the atomic **or** molecular weight of the radiating atom **or** molecule. Since T is the absolute temperature, the line width may be used to calculate the temperature of the solar atmosphere at whatever height above the solar surface the observer desires.

Laboratory studies using an interferometer have shown that the breadth of the krypton line (λ = 5570Å) at liquid air temperatures is 0.006Å, practically all of which can be ascribed to the Doppler effects resulting fromthermal agitation at that temperature (80°K).

(b) Pressure **or** "Statistical" Broadening - The light from a single unresolved spectral line may be made to produce interference fringes when the path difference is as much as several hundred thousand wavelengths. The implication of collisions between atoms is verified the higher the pressure of the gas for a given temperature. An increase in temperature for a given density also increases both the collision rate and the pressure. Thus, in general, at higher gas pressure, there is a broadening of the spectral line due to the Doppler effect and a broadening due to the increasing frequency of phase changes resulting from collision. Michelson found that below a pressure of the order of a millimeter of mercury, the breadth of the hydrogen line λ = 6563Å is almost entirely due to the Doppler effect; but that at higher pressures the line becomes considerably broader.

(c) Natural Line Breadth - The minimum width of spectral lines is limited by the Heisenberg uncertainty principle, $\Delta E \Delta t = \hbar$ In the visible region, the natural line breadth is much less than 0.001Å and is not detectable. It is easily observable in the case of X-rays.

(d) Stark Effect - Spectra lines are broadened by small electric fields which are not large enough to produce an observable splitting of the line.

(e) Turbulence - Motions of the solar gases, due to turbulence, that have a component in the line of sight increase the thermal Doppler width by a factor of $\sqrt{2}$. Turbulence does not depend on molecular weight, while Doppler Broadening does.

(f) Radiation Damping - Classically, radiation damping is explained as due to the accelerated motion of the electron and loss of energy by radiation. The classical damping constant, γ , is given as

$$\gamma = \frac{0.22}{\lambda^2} \simeq 10^8 \qquad \text{(S1-11)}$$

In the quantum theory of radiation, the damping is found to be associated with the transition probabilities and the finite width of the atomic levels.

(g) Einstein Shift - Relativity predicts that radiation quanta leaving a massive source of radius R , such as the sun, would be retarded by the gravitational attraction and hence lose energy. The loss of energy would cause a decrease in frequency resulting in a shift toward the red end of the spectrum. The observed frequency is given by

$$f' = f\left(1 - \frac{GM}{c^2 R}\right) \qquad \text{(S1-12)}$$

where G is the gravitational constant and M is the mass of the source body. Applying the necessary values to GM/c^2R generates a spectral line shift of

$$\frac{\Delta\lambda}{\lambda} = 2.02 \times 10^{-6} \qquad \text{(S1-13)}$$

St. John[13] (1928) made a comprehensive study based on 1537 spectral lines at the center and 133 at the edge of the sun, with a probable error for a single source line being ± 0.0008A. In general, St. John's results confirm the predicted relativistic shifts given in the following table.

TABLE II(S1)

PREDICTED EINSTEIN SHIFTS

Wavelength (Å)	3800	4250	4725	5675	6600
Displacement (Å)	0.0008	0.009	0.010	0.012	0.012

VII. REFRACTIVE INDEX

As early as 1911, A. Cotton[15], postulated that solar atmospheric turbulence and time varying refractive indices could induce spectral line shifts, thus generating anomalous radial rotational velocities.

Recent calculations[16], have deduced the index of refraction of the solar atmosphere to be $n = 1 \pm 10^{-6}$, and therefore has a negligible effect.

APPENDIX S-1A

DERIVATION OF THE EQUATIONS OF MOTION

For a spherically symmetrical distribution of matter in gravitational equilibrium:

$$\begin{aligned} M(R) &= \int \rho \, dV(R) \\ dM(R) &= \rho \, dV(R) \\ dM(R) &= 4\pi R^2 \rho \, dR \end{aligned} \qquad \text{(S1-14)}$$

where $M(R)$ is the mass included in a sphere of radius R, ρ is the density, and V is the volume. Consider an infinitesimal cylindrical shell at a distance r from the center with a thickness dr and a unit cross-section at right angles to r.

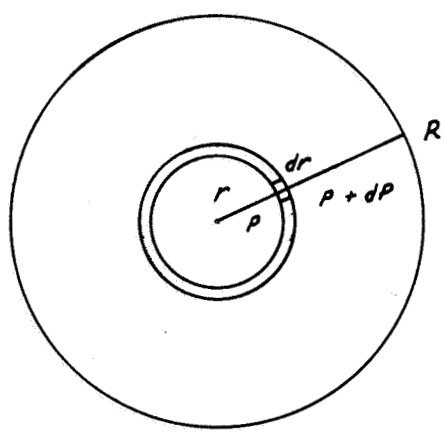


Figure 4 (S1) ILLUSTRATION OF EQUATIONS

P is the pressure at r and dP is the increment in P going from r to $r + dr$. The difference in pressure represents a force, $-dP$, acting in the direction of increasing r. For gravitational and hydrostatic

equilibrium, the outward force must be balanced by the gravitational attraction and centripetal force, mv^2/r, due to rotation to which the element of mass is subjected. The force of attraction between $M(r)$ and ρdr is

$$\frac{GM(r)}{r^2}\rho dr \tag{S1-15}$$

Additionally, the attraction due to the material outside r is zero, and the mass $M(r)$ is assumed as a concentrated point mass at the center. Then

$$dP_r = -\frac{GM(R)}{R^2}\cos\phi\,\rho dr + \alpha_r \rho dr \tag{S1-16}$$

where $a_r = r\omega^2 = v^2/r$, $r = R\cos\phi$

and ϕ is the solar latitude.

Then

$$\frac{dP_r}{dr} = (g_r + \alpha_r)\rho \tag{S1-17}$$

letting

$$g_r = \frac{-GM(R)}{R^2}\cos\phi = \text{component of gravity in the } r\text{-direction} \tag{S1-18}$$

results in

$$\frac{dP_r}{dr} = (g_r + r\omega^2)\rho \tag{S1-19}$$

Finally, putting (S1-14) and (S1-18) in (S1-19) and assuming $d\omega = 0$, the fundamental equation of equilibrium becomes:

$$\frac{1}{r^2}\frac{d}{dr}\left(\frac{r^2}{\rho}\frac{dP}{dr}\right) - 3\omega^2 = -4\pi G\rho\cos\phi \qquad \text{(S1-20)}$$

where. ϕ is the solar latitude.

The equation of motion in the z-direction is obtained in the same manner. Non-inclusion of the angular velocity term gives:

$$dP_z = \frac{GM(R)}{R^2}\rho\sin\phi\, dz = g_z\, dz \qquad \text{(S1-20)}$$

whose fundamental equation is

$$\frac{1}{z^2}\frac{d}{dz}\left(\frac{z^2}{\rho}\frac{dP_z}{dz}\right) = -4\pi G\rho\sin\phi \qquad \text{(S1-21)}$$

where

$$z = R\sin\phi .$$

APPENDIX S-1B

DERIVATION OF ENERGY GENERATION EQUATION

To obtain the energy generation rate equation we use Von Zeipel's theorem which assumes hydrostatic and radiative equilibrium for a rotating rigid body. Two consequences of theseassumptions are: (1) a rotating star is brighter at the poles than at the equator; (2) the energy generation law must be of the form

$$\mathcal{E} = const. \left(1 - \frac{\omega^2}{2\pi G\rho}\right)$$

Taking the origin to be the center of mass, the effect of rotation is to increase the gravitational potential, then

$$W = V + T_{ROT} = V + \frac{1}{2}\omega^2(x^2 + y^2) \qquad \text{(S1-22)}$$

where W is the energy/unit mass and ω is the angular velocity of rotation where the Z-axis is the axis of rotation.

For hydrostatic equilibrium

$$grad\ P = \rho\ grad\ W \qquad \text{(S1-23)}$$

since

$$\nabla^2 V = -4\pi G\rho$$

is satisfied, we may resubstitute

$$\mathbf{V} \cdot \left(\frac{1}{\rho}\nabla P\right) = -4\pi G\rho + 2\omega^2 \qquad \text{(S1-24)}$$

Radiative equilibrium requires

$$\nabla T^4 = -\frac{3}{ac} K\rho F_r \qquad \text{(S1-25)}$$

where F_r is the radiative flux, T the temperature. Now

$$P_r = \frac{1}{3} aT^4 \qquad \text{(S1-26)}$$

P_r is the pressure due to monochromatic radiation. Then

$$F_r = -\frac{c}{K\rho} \nabla P_r \qquad \text{(S1-27)}$$

and

$$\nabla \cdot F_r = -\frac{c}{K\rho} \nabla^2 P_r = \rho\varepsilon \qquad \text{(S1-28)}$$

since $\nabla \cdot F_r = \rho E$ is the energy produced in a shell of mass ρ and ε is the energy produced per gram per second. Combining

$$-\frac{\varepsilon\rho}{c} = \nabla\left(\frac{1}{K\rho} \nabla P_r\right) \qquad \text{(S1-29)}$$

The physical picture implies

$$\nabla P_r \propto \nabla P$$

Then:

$$div\left(\frac{1}{\rho} grad\, P\right) = -4\pi G\rho + 2\omega^2 \qquad \text{(S1-30)}$$

$$= -\frac{\varepsilon\rho}{c}$$

and

$$\varepsilon = const. \left(1 - \frac{\omega^2}{2\omega G\rho}\right) \qquad \text{(S1-31)}$$

This condition is physically unreasonable since it implies that the rate at which energy is produced per gram per second is dependent on the angular velocity, and hence one concludes that stars do not rotate as rigid bodies.

VIII. REFERENCES

1. Starr, V. P. and Gilman, P. A., Energetics of the Solar Rotation, Ap. J., 141, 1964, pp. 1119-1125.

2. Richardson, R. S. and Schwarzschild, M., Convegno Volta, Accademia Nazionale dei Lincei, 11, 1953, p. 28.

3. Ferraro, V. A. C., M. N., 97, 1937, p. 458; Sweet, P. A. M. N., 109, 1949, p. 507.

4. Walen, C., Ark, f. Mat., Astro. Och. Rys., 33A, 18, 1946.

5. Cowling, T. G., The Solar System, Vol. 1; The Sun, ed. G.P. Kuiper Chicago University of Chicago Press, Chapt. viii.

6. Ward, F., Astro. Phy. J., 2, 1965, p. 441.

7. Ward, F., J. Pure and Appl. Geophys., 58, 61, 1964.

8. Newton, H.W., and Nunn. M.L., M.N., 111, 413, 1951.

9. Aller, L. H., Astrophysics, The Atmosphere of the Sun and Stars, New York: Ronald Press Co., Chapt. ix, 1963.

10. Jenkins, F. A., White H. E., Fundamentals of Optics, New York: McGraw-Hill Book Co., 1957, p. 403.

11. Sommerfeld, A., Optics, New York; Academic Press, Inc., 1964, p. 69.

12. Richtmyer, F. K. and Kennard, E. H., Introduction to Modern Physics, New York; McGraw-Hill Book Co., 1947, p. 396.

13. St. John, C. E., Ap. J., 67, 195, 1928.

14. Blanco, V. M. and McCuskey, S. W., Basic Physics in the Solar System, Reading, Mass.: Addison-Wesley Publishing Co., 1961, p. 90.

15. Cotton, A., Astro. Phy. J., 33, 375, 1911.

16. Minnaert, M., The Solar System, Vol. 1: The Sun, ed. G. P. Kuiper, Chicago: University of Chicago Press, 1953, p. 130.

17. DeJager, C., Structure and Dynamics of the Solar Atmosphere, Vol. LII, Encyclopedia of Physics, ed., S. Flugge, Heidelberg, Germany: Springer-Verlog OGH), 1959, p. 339.

APPENDIX S-2

SOLAR GRANULATION

By Robert O. Breault

This appendix describes the phenomenon of solar granulation. It includes a discussion of the observational results and difficulties as well as the relation of granulation to other phenomena.

I. INTRODUCTION

The surface of the sun exhibits a random graininess structure termed granulation. The physical mechanisms which generate the convective appearance of granules are unknown and the present state-of-the-art knowledge may be considered as an ensemble of empirical inferences. To propound a photospheric fine structure model demands detailed knowledge of the solar surface, which, because of the optical properties of earth-based instrumentation and atmospheric constraints, appears unattainable.

Surface phenomena manifested in the form of vertical and horizontal velocity fields appear to have independent energetic driving forces all existing in an environment of varying magnetic field magnitude. Physically, the characteristics of chromospheric plages and spicules perpetuate as if they are by-products of photospheric interactions, yet there is no evidence for a direct or indirect connecting supposition.

II. GRANULATION

High resolution photographs of the solar surface show a rice grain pattern of matter similar to convection cells obtainable under laboratory conditions. Problems concerned with convection are approached using a scheme devised by Bénard and expanded by Rayleigh. Experimentally, it was shown that if a thin layer of liquid is subjected to a uniform temperature T on its lower surface and a lower temperature on its upper surface, and if the temperature difference, ΔT, exceeds a certain critical value,

convection occurs in a definite cellular pattern, where the cells take the form of regular polygons. The criterion for convection to occur is given as

$$\Delta T > R \frac{K\nu}{g\alpha} d^{-4} \qquad \text{(S2-1)}$$

where g is the acceleration of gravity, d is the thickness of the layer, and α, K and ν are, respectively, the coefficients of volume expansion, thermal conductivity, and kinematic viscosity. R is a non-dimensional quantity called the Rayleigh number[1]. The applicability of Bénard's work to the photosphere is limited and gives only a first order of approximation, but the similarity in the physical structure of granular cells to laboratory convection patterns has led to the view that granules represent convection cells.

Granules have been identified within 10" of arc of the solar limb. The determination of the distance from the limb at which granular patterns finally disappear provides an estimate of the height of the top of the stable convection zone from which the granules are believed to have originated. Sub-photospheric strata are assumed to be stable convection zones terminating with convective instability on the surface. Immediately a conflict arises. Convective currents require a different temperature gradient than that needed for radiative transport of energy. Although photospheric granulation suggests that convective currents exist in the upper layers of the sun, the temperature gradient demanded by limb-darkening observations is more nearly in accord with that appropriate to the transport of energy by radiation[2]. The empirical paradox is unexplained.

Granules cover the entire disk of the sun, and vary both in size and number with the solar cycle. From high-quality Stratoscope photographs, it was determined that the fading[3] of the granular pattern occurs primarily at latitudes of $60" < \phi < 78"$ and that the fading granulation is replaced by a larger, low-contrast pattern observable out to $\phi \simeq 87°$. Additionally, inhomogeneities in the granulation pattern consisting of low-contrast regions that persist longer than the average lifetime of the granules themselves and small-scale regions which appear to evolve with the normal pattern have

been observed. The photographs used were taken above the terrestrial atmosphere and cover a wavelength band of 800Å centered at 5450Å. Photospheric granulation in the central region of the solar disk consists of cellular patterns of bright granules superimposed on a darker background. The majority of the granules have diameters that range in size from 300 km, which is the diffraction limit of resolution for a 12-inch telescope placed external to the terrestrial atmosphere, to about 1800 km, the mean diameter being about 700 km. In shape, the granules simulate irregular polygons, some of which appear elongated, and although they are not indicative of Bénard cells, they are not entirely random phenomena either. The granular cells are separated by narrow dark lanes whose apparent widths often do not exceed a few tenths of a second of arc[4]. Schwarzschild has suggested that granules represent a "state of non-stationary convection in which a range of cell sizes exist, the cells having irregular shapes caused by rising columns of material with the dark narrow lanes representing sinking material". Thus, the solar granules would lie between a domain of stable convection and classical turbulence.

As a guiding parameter, independent of photographic contrast, instrumental profile and terrestrial atmosphere, workers in the field have employed measurements of the mean cell size, defined as the average distance between the centers of adjacent granules, for characterizing the scale of the pattern. Figure 1(S2) is a histogram of mean cell size. The granulation appears to have a well-defined and rather narrow distribution of sizes with 70% of the values lying between 2".2 and 3".8 of arc. The mean value is 2".9.

Recent work[5] has suggested a mean value of 2".5 as a more appropriate representation.

Observations away from the central disk show an increasing foreshortening of the granulation pattern toward the limb until finally the granules fade and become unresolvable about 10" of arc from the limb. Although the granules are not distinct near the edge of the disk, we see instead, large bright areas (faculae) appearing at somewhat higher levels[6].

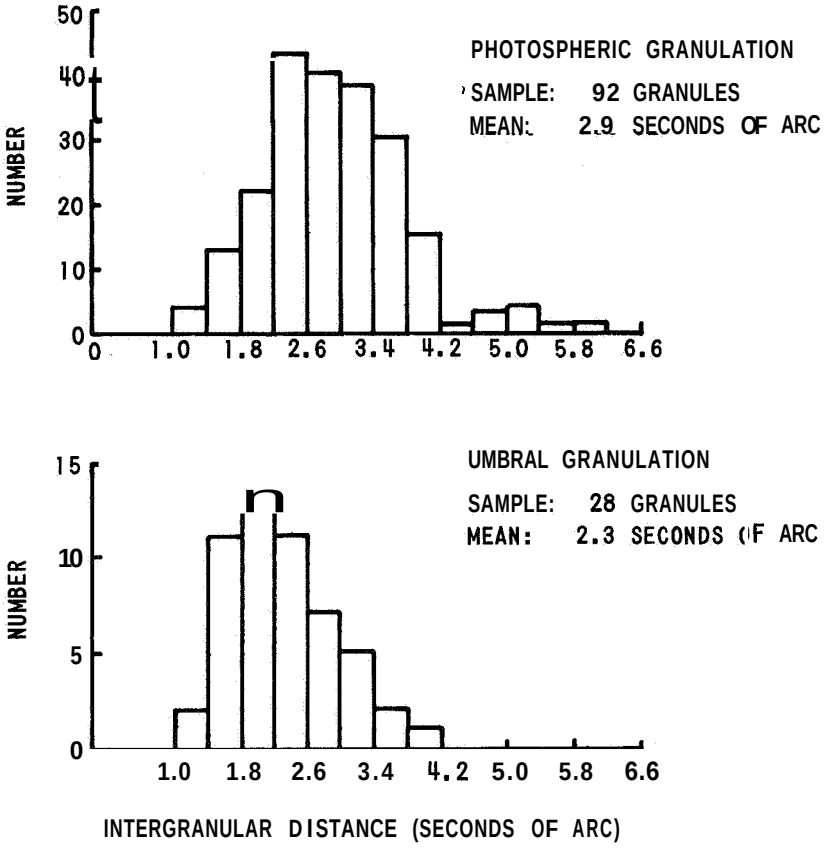


Figure I (S2) CELL SIZES OF GRANULATION PATTERNS. THE HISTOGRAMS GIVE THE DISTRIBUTIONS OF THE DISTANCES BETWEEN THE CENTRES OF ADJACENT GRANULES (CORRECTED FOR FORESHORTENING). (REF. 4)

The faculae are inseparable companions of sunspots and a physical connection between granules, sunspots and faculae has long been sought. To date, no empirical evidence has been found.

Two additional characteristics of granules are their lifetimes and brightness fluctuation. The accepted average length of time between the formation and decay of a granule is 10 min. Researchers have found that once a granule is formed, its diameter increases until it reaches about 2" of arc, whereupon it disintegrates into several small granular fragments that disappear at the same spot as the granule originated. The lifetime of

granules found in sunspots suggests some support for a strong inhibiting effect of a magnetic field on the establishment of convection. Increases in lifetime to 30 min for umbral granules and to several hours for penumbral filaments have been detected. In addition, umbral granules appear to be closer together than photospheric granules and show no change in brightness or shape during their lifetime [Figure 1(S2)].

The centers of the granules are hotter than the intergranular regions by 50°K to 100°K but the intensities of the individual granules fluctuate along with changing size and shape. A recent study of 140 granules resulted in the following statistics: 57% showed no change in brightness, size or shape, 8.270 showed a brightness increase, 8.270 showed a brightness decrease, 3% showed both an increase and decrease in brightness, 11.5% gave an increase in size, 5.7% gave a decrease in size, and 12.270 showed a change in shape.

Many estimates have been made of the magnitude of brightness variations in granular patterns. Schwarzschild found a root-mean-square brightness variation of ±93°K between granular cells and intergranular spacing. The contrast of the granulation is as yet undetermined for two reasons: (1) the value of the contrast transmission function at the spatial frequency corresponding to 0".5 (width of dark lanes) was found to be 0.05, (2) conventional contrast transmission function refers to unidimensional brightness distributions, whereas the granulation pattern constitutes a two-dimensional brightness distribution.

III. VELOCITY FIELDS

The spectra of granules show Doppler shifts indicating that the granules themselves are columns of hotter gases rising from sub-photospheric layers. As the gas columns reach the top of the photosphere they spread out and sink down. The darker intergranular lanes are considered to be the cooler gases sinking back into the photosphere. Displacements in the spectral lines give a mean random velocity of 0.37 km/sec for cells of the order of 2" or 1500 km in diameter[7]. The employed spectral lines

(Fraunhofer) originate in the photosphere, above the granulation, and thus there appears to be an "overshoot" from the granulation cells up into the more tenuous photosphere.

Investigators[8] have found a distinct correlation between brightness fluctuations and vertical velocities. The characteristic "cell size" of the vertical velocities appear to increase with height above the solar surface, growing from approximately 1700 km to 3500 km. The root-mean-square vertical velocity was found to be constant at 0.4 km/sec during this growth period. Also, the vertical velocities exhibit a repetitive time correlation, with a period $T = 296 \pm 3$ sec. Local vertical velocities in the solar atmosphere are considered non-random due to their striking quasi-oscillatory small-scale motions is the buoyancy of the hot granules. Additionally, the small scale brightness fluctuations in the lower chromosphere appear to exhibit the same periodic time correlation as the Doppler velocities, thus implying that the oscillatory motions are probably responsible for these brightness fluctuations.

On the basis of the aforementioned empirical evidence, a model which gives a plausible explanation of the mechanism by which energy is transferred from the granulating layer into the chromosphere has been proposed[9]: "Gravitational energy in the form of buoyancy of hot granules is converted into kinetic energy in the lower and middle photosphere, as evidenced by the brightness-velocity correlation, and thence, through the elastic properties of the atmosphere, into an oscillatory wave of definite frequency. This wave, propagating upward, deposits its energy in the chromosphere. On the basis of the numerical results so far obtained, it seems quite possible that the oscillatory motions contribute a major part of the energy which heats the chromosphere in non-active areas."

IV. HORIZONTAL VELOCITY FIELDS

In 1960, Leighton discovered systematic horizontal currents on the surface of the sun. He called this horizontal field motion "super-granulation" analogous to photospheric granules. The super granules are considered as a large-scale counterpart to photospheric granulation, having

cell diameters of 15,000 kilometers and lifetimes on the order of several hours. These super cells can only be detected spectroscopically and are not visible in white light. The motion of moving material within each cell indicates a central source with all material flowing toward the cellular periphery at a velocity of about 0.5 km/sec. The super cells are related geometrically to structures observed in the chromosphere. The suggestion is that the cellular patterns and chromospheric networks are surface manifestations of a super system of convection current originating deep in the convective layers. There exists strong corroboration between independent researchers, that large scale fields cover the entire solar surface and not just the equatorial latitudes.

The overall physical picture of photospheric motion suggests that the velocity field is composed of two distinct parts: the large scale cellular pattern of long-level horizontal motions and the small scale granular pattern of quasi-oscillatory vertical motion. The true nature of the two fields and the energy agencies from which they are supplied are unknown.

V. MAGNETIC FIELDS

There seems to be little correlation between velocity and magnetic fields except that maximum field strength tends to occur along lines of minimum velocity. To date, the specific magnetic character of granulation is unknown because of inadequate resolution generated from telescopic imperfections and seeing conditions.

In the H_α chromosphere, there is a strong variation in the velocity field with altitude. At high elevations, a vertical velocity field with short lifetimes (30 sec) are observed. There is no strong indication of association with local magnetic fields, while at lower elevation over an existing quiescent surface, material motion is predominantly downward through a network of funnels indicating strong participation of magnetic field. The lifetime of the velocity pattern is on the order of several hours, although significant changes do occur within a few minutes.

A point of interest is that according to the general theory of relativity, spectral lines produced in a quiescent, homogeneous, solar atmosphere should show a red shift given by

$$\frac{\Delta\lambda}{\lambda} = \frac{GM}{RC^2} \approx 0.3 \text{ km/sec} \qquad \text{(S2-2)}$$

A recent magnetic scanning method developed by J. E. Blamont enables one to observe a line profile with a resolution of 10^{-3} Å. At the center of the solar disk, he found a red shift which corresponded to the predicted value.

VI. PLAGES AND SPICULES

When the sun is viewed in monochromatic light, high intensity chromospheric areas, in the regions of sunspots, appear. These areas, called plages, are on occasion found in higher-than-average magnetic field regions with no visual sunspot activity. The chromospheric plages are locations where the atoms of an element are changing their states of ionization or excitation at optical wavelengths. Calcium and hydrogen plages appear to have a spatial and temporal relationship since they usually occur in approximately the same region at the same time. When the plage regions emit radiation at several frequencies and become visible in "white-light", i. e., in the direct image of the sun, they are called faculae.

Investigators have reported[10] that Ca II and H_α plages show a close association with regions of increased magnetic field strength. Plage regions have been associated with Type VI (Coronal) prominences and solar flare activity. Additionally, X-ray emission sources are found in local coronal areas overlying the calcium plages. These same areas produce a slowly varying radio component and require temperatures of 0. 6 to 1. 5 million degrees to explain the radio emission. A correlation between plage brightness and coronal temperature appears positive[11], i. e., the brighter the plage the higher the temperature. From plage and direct magnetic field measurements one would think that magnetic fields occur where plages appear, but fine detail is lost in large field resolution and the relationship remains speculative.

Another chromospheric anomaly, with an apparent physical connection to photospheric granules, are spicules. When the chromosphere is viewed near the limb, many jet-like spikes of gas are seen rising vertically through it. These spicules, when viewed in the light of hydrogen consist of gas jets moving upwards into the chromosphere with speeds near 20 km/sec, and last for only a few minutes. The "spiky" structure commences about 4000 km above the photospheric surface and extends upward into the corona, thus, spicules appear as cool elements in the hot (1×10^6°K) outer atmosphere. Based on observations of emission-line profile broadening, recent measurements[12] indicate a spicule temperature of 20, 000°K.

VII. CONCLUSIONS

Granulation is the primary feature of solar topography in the photosphere. Its observation requires high resolution. Since angular resolution in space is a function of the optical system employed, the tacit advantage of a space probe's ability to achieve a specified position with respect to the sun is to gain a higher diffraction-limited linear object resolution and hence a detailed fine structure analysis of the granular surface. An increased object resolution would necessarily expedite the accretion of solar topographical information, exemplified by the following proposed quasi-functional relationships: limb-darkening observations and convective surface instability; faculae and sunspots; granular lifetime and granular contrast; spectra analysis of sub-photospheric layers and the element abundance; vertical and horizontal velocity fields; chromospheric spicules and photospheric granules; plages, sunspots and their interconnection with magnetic fields.

(The reader is referred to Appendix S-6A where some considerations and calculations are presented concerning the difficulty of obtaining stereoscopic photography of solar granulation.)

Our knowledge of these relationships would be significantly increased if they could be studied from information gathered outside of the terrestrial atmosphere, as demonstrated by successful Stratoscope experiments.

Granular structure, as a physical phenomenon in itself is of limited value, but the interaction of the photospheric granules with other surface phenomena is of primary interest, since the exact nature of these mechanisms must be known prior to the construction of a successful solar model.

VIII. REFERENCES

1. Aller, L. H., "The Atmosphere of the Sun and Stars", Astrophysics, Ronald Press Company, New York, 1963, p. 451.

2. Ibid, p. 233.

3. Edmonds, F.N., Ap. J., 131, 1960, p. 57.

4. Bray, R. J. and Loughhead, R. E., Sunspots, John Wiley & Sons, Inc., New York, 1965, p. 55.

5. Macris, C. J. and Banos, G. J., "Mean Distance Between Photospheric Granules and Its Change with the Solar Activity", Men. Nat. Obs. Atheres: Serial 1, No. 8, 1961.

6. Baker, R., Astronomy, 8th Ed., D. Van Nostrand Company, New York, 1964, p. 290.

7. Aller, L.H., Op. Cit., p. 455.

8. Leighton, R. B., Noyes, R. W., and Simon, G. W., Ap. J., 135, 1962, p. 474.

9. Ibid, p. 495.

10. Athay, R. G., Thomas, R. N., Physics of the Solar Chromosphere, Interscience Publishers, Inc., New York, 1961, p. 47.

11. Kawabata, K., Publ. Astron. Soc. Japan, 12, 1960, p. 513.

12. Zirker, J. B., Ap. J., 135, 1962, p. 515.

APPENDIX S-3

SOLAR FLARES

By D. A. Richards

A brief description of solar flares is presented. The importance of these phenomena is explained to establish motivation for their study. The correlations with other solar phenomena are described. These correlations are studied in an attempt to attain prediction capability. Some ideas are presented for space observations that might assist the solution of the solar flare problem.

I. INTRODUCTION

During an investigation of the sun, one necessarily must consider the various anomalous phenomena of its topography. The most important of these as far as man is concerned is probably solar flares. Their apparent effects on the earth's weather and communications and the possible danger they may present to space travelers has resulted in the great amount of research currently being conducted and promoted in this area of solar study. The following exposition is intended to present a brief description of flares to include their correlations with other solar phenomena and some ideas for space observations to advance prediction capability.

II. SOLAR FLARES

A solar flare is a disturbance in the atmosphere of the sun that is apparent as a region of exceptional brightness which develops very suddenly in a plage area of the chromosphere as generally observed in the monochromatic light of the hydrogen alpha line. There are several different types of solar flares, as indicated by their appearance and behavior. These factors are taken into consideration in the system commonly used to classify flares according to their "importance". A table of characteristics is given below.

TABLE I(S3)

SOLAR FLARE CHARACTERISTICS

Importance	Average Duration	Range of Duration	Area in Millionths of Solar Disk	Relative Frequency
1	20 min.	4-40 min.	100-250	72%
2	30 rnin	10-90 min.	250-600	23%
3	60 min	20-150 min.	600-1200	4%
3t	180 min.	1-7 hrs.	>1200	1%

Although much observational data on the effects of flares is available, their origin and the physical processes which produce these effects are, as yet, not fully understood. Most of the answers offered are primarily speculations that demand further corroboration from observations and theory.

Flares derive their major significance from the high correlation of greater solar cosmic radiation emission and the occurrence of superflares (3t importance). Although solar protons are the predominant radiation resulting from solar flares, alpha particles, heavier nuclei and X-rays are also present. These radiations represent a potential hazard to man during spaceflight and also produce many important effects on the earth due to their interactions with the magnetosphere and atmosphere. These interactions produce magnetic storms and consequent radio disturbances as well as auroral phenomena. The threat to man in space and the disruption of communications are the primary reasons for studying flares and their attendant eruptions.

Another major reason for the interest in the cosmic-ray events of solar flares is the opportunity such phenomena offer in the study of interplanetary magnetic fields. The particles produced in these events continue to rain upon the earth for a period of hours after the disappearance of the flare from which they arose. Apparently the particles cannot escape directly from the solar system, but are trapped temporarily within it. Those

particles observed immediately after the flare often appear to approach the earth from a fairly well defined direction; toward the end of the event the particles appear to come from all directions. This is explained by describing the early particles as those which come directly from the sun and the late particles as those which have undergone scatterings before they reach the point of detection. Even the early particles, of course, do not usually come precisely from the direction of the sun, because the magnetic fields present in interplanetary space bend their trajectories.

If we assume for the moment that solar cosmic rays are exclusively a phenomenon of major solar flares, the prediction problem does not appear too difficult. Flare incidence correlates moderately well with sunspot number, and particularly with the morphology of individual active centers. The penumbral area of spots has been investigated as to its usefulness for flare prediction. Of course, more comprehensive prediction criteria than a single parameter should be used. Solar-cosmic-ray emission is mainly a large-flare effect. Sunspot area and morphology should be supplemented with east-limb coronal intensities, the magnetic history of the spot group, the incidence of small and large flares in the region, the size and complexity of the associated plages, and especially the radio frequency emission from the sun.

Cosmic-ray flares exhibit a definitely asymmetric distribution on the solar disk. (A preference for the northwest quadrant has been noted.) It might be of value to establish flare patrols (by probe or satellite) on the other side of the sun to obtain additional information about preferential formation. Since transmission cannot traverse the sun, such a system would require a design which allows for storage and choice of transmission time. Because flares are generally not predictable and of short duration, solar flare patrols established in space are very apt for adaptive system use. The majority of the information collectable during such a mission would not be of value to the objectives of understanding and predicting flares or their cosmic ray interactions with the interplanetary magnetic fields.

During periods of no flare activity, the probe or satellite could monitor the solar wind, solar magnetic field (relatively undisturbed) or the galactic cosmic radiation. An adaptive system could sample readings from the instrumentation at intervals determined by the rate of change of the reading, store the readings and transmit when some predetermined amount of information has been collected. This would provide a saving in storage requirements for the monitoring of relatively constant fields and particle fluxes.

The most pressing problem is to discover preflare characteristics of active centers that mark them as likely to emit cosmic radiation when they produce great flares. What is needed is a detailed time sequence of the magnetic fields - their strength, polarity, direction, and rate of change with time - in the regions of near-zero longitudinal field intensity where flares have a tendency to break out. Continuous coverage of these regions with a battery of sensors could be achieved for maximum results by recording their outputs in storage and erasing after some predetermined delay, unless a transmit command was generated by the onset of a flare.

III. CONCLUSIONS

It would require a more detailed study of the observational data concerning solar flares and solar cosmic radiation to determine what instrumentation systems could best be employed to measure them, their effects and clues to their prediction. The wide range of values encountered in energetic particles and magnetic fields are suitable for automatic range selection circuitry. The advantages of adaptive systems might be found from a more careful analysis of the vast amount of redundant information which could be collected from a system designed to investigate solar flares.

REFERENCES

Abetti, G., "Solar Research", The MacMillan Co., New York, 1963.

Berkner, L. V. and H. Odishaw, Ed., "Science in Space", McGraw-Hill, New York, 1961.

3. Kuiper, G. P., Ed., "The Sun", The University of Chicago Press, Chicago, 1953.

4. LeGalley, D. P. and A. Rosen, Ed., "Space Physics", John Wiley & Sons, New York, 1964.

5. Parker, E. N., "Interplanetary Dynamical Processes", Interscience Publishers, New York, 1963.

6. Rossi, B., "Cosmic Rays", McGraw-Hill, New York, 1964.

7. Smith, H. J. and E. P. Smith, "Solar Flares", The MacMillan Company, New York, 1963.

APPENDIX S-4

SOLAR ATMOSPHERIC MODELING

By Robert O. Breault

This appendix is concerned with the construction of solar atmospheric models. It describes the important concepts, theories and parameters necessary to such an endeavor, as well as the general procedure to be followed.

I. INTRODUCTION

A major purpose for acquiring information pertinent to any astrophysical body is model construction. A solar photospheric model should be an accurate representation, based on the observable features in, on, or near the visible layer of the Sun's surface, in terms of the Sun's temperature, pressure and chemical cornposition.

The mathematical formulation resulting from the experimental and theoretical considerations which represent the physical and chemical variables of the Sun will comprise the text of this work.

The theory of atmospheric models demands a determination of the parameters involved in a model from the observed spectral features. In principle, from the fundamental assumptions of physical structure, radiative equilibrium and known elemental abundances, all information needed in the construction of a model solar atmosphere is available. Realistically, geo-atmospheric constituents severely hamper gathering information about celestial bodies. Additionally, broad underlying assumptions and difficult or unattainable measurements result in approximations, which in turn generate deviations from the desired result. The extent of this departure, if known, is reported in each appropriate subsection.

II. ORGANIZATION

The following sections appear as independent units. This has been done for ease of handling. The connections among them become apparent, however, as they are integrated into the total model in Section VII.

The third section explores the restrictions and limitations of earth-based observations. Section IV gives the mechanics necessary for postulating a plane layer atmosphere and the associated assumptions which are made that express the physical picture of a star in equilibrium. The object of this section is to establish the requirements for hydrostatic equilibrium and to determine the parameter gradients required for various modes of energy transport. Section V assumes a photosphere in radiative equilibrium and attempts to define and explain the parameters which must be known in order to maintain the validity of this assumption. Section VI defines the theoretical and empirical information necessary for calculating the hydrogen-to-metal ratio.

Section VII combines the assumptions, postulations, data and theory of Sections IV, V and VI in an attempt to explain, step-by-step, what the process of model building entails. This section additionally reproduces several suggested photospheric models. The final section contains conclusions.

Figure 1(S4) is an illustration schematically expressing the dependence of the assumptions on their parameters and the interrelationship of the assumptions with each other.

It should be mentioned that the concepts utilized in the construction of a photospheric model, in the ensuing work, have been oversimplified and condensed in order to insure continuity of thought for a wide group of readers. For a comprehensive treatment of solar models, the reader is referred specifically to References 1, 2 and 6.

III. INFORMATION DISCRIMINATION

Astrophysical phenomena reveal their physical nature and/or the associated chemical composition of their supporting material through emission, absorption, scattering or reflectance of electromagnetic radiation.

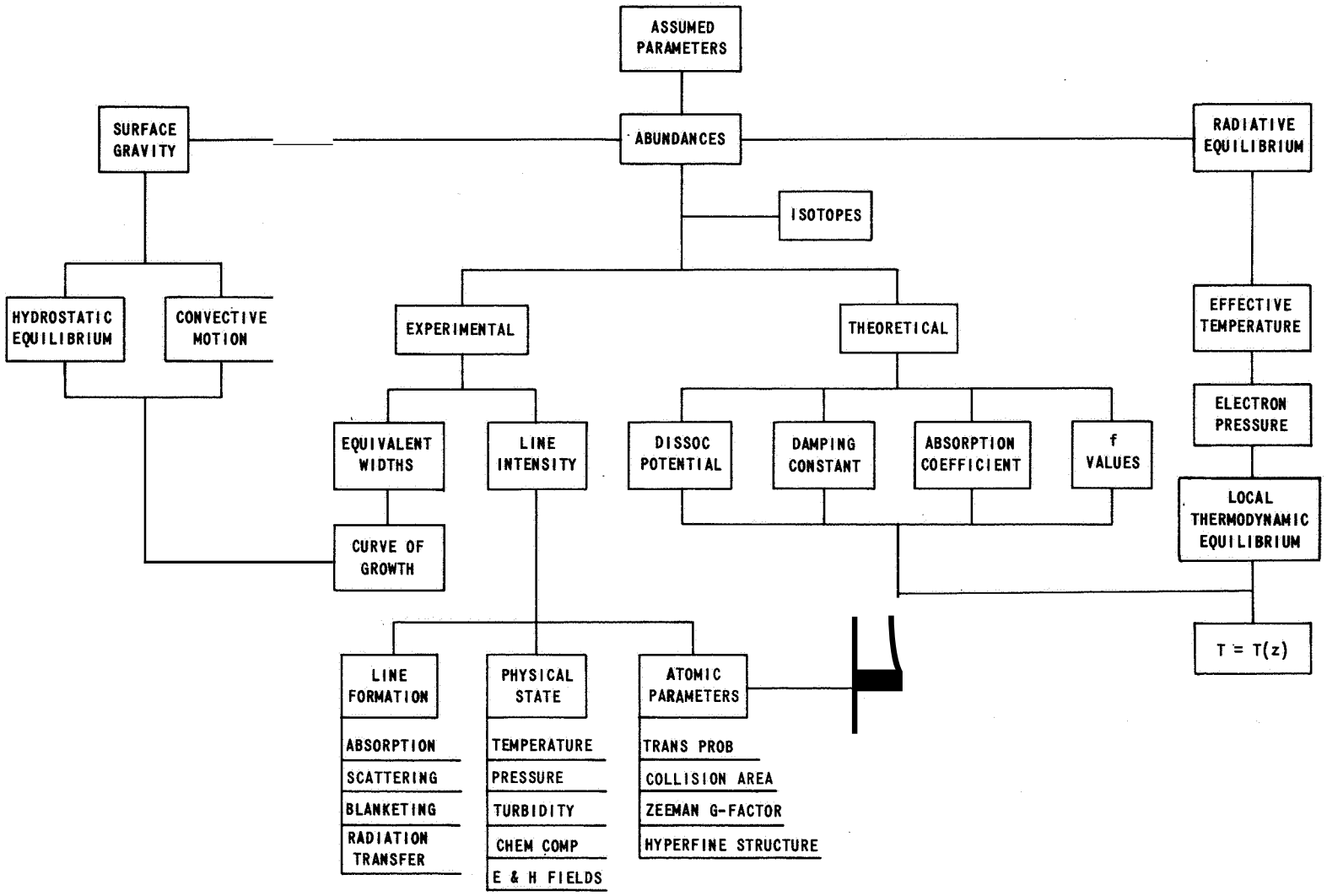


Figure 1 (S4) ASSUMPTIONS AND THEIR PARAMETERS

The continuous and darkline spectra observed are the principal sources upon which interpretation may be based. Unfortunately, information gathering devices are unable to exploit the major portion of the radiation because of partial or complete atmospheric absorption.

Those zones where there is a minimum of atmospheric absorption, called windows, occur in two areas, the optical region and the radio-frequency region [Fig. 2(S4)]. From the top band of the figure it can be seen that except for the realm, $20\text{Å} - 912\text{Å}$, caused by interstellar media consisting of hydrogen and helium, the entire electromagnetic spectrum from X-rays to radio-rays may be studied from above the atmosphere.

The lower half of Fig. 3(S4) represents an enlarged-portion of the optical window. In the infrared there exists strong absorption bands of O_3, CO_2 and H_2O, implying that the study of a planetary atmosphere, whose average temperature is $T = 300°K$, would be severely hampered. From the upper half of Fig. 3(S4) it can be seen that the bulk of the solar radiation falls in the optical region and is relatively free from absorption bands. It is in this small region, from $\log \lambda$ (Å) = 3.5 (3100;) to $\log \lambda$ (Å) = 3.85 (7100;) that the physical features of a star can be examined from the earth.

IV. STELLAR STRUCTURE

By assuming that the depth of the photosphere is small compared to the radius of the Sun, the curvature may be neglected and the solar surface treated as a plane layer. Under these conditions the acceleration due to gravity, $g = 2.74 \times 10^4$ cm/sec^2, for the Sun may be assumed constant. A first order approximation results when the atmospheric layer is treated as composed of solar gases that are homogeneous over the whole depth of the photosphere and are maintained at a constant temperature. A condition that must be met if the above assumptions are to retain their validity is the establishment of hydrostatic equilibrium. Additionally, one must also consider the occurrence of convective layers in turbulent motion. Subsections 1 and 2 express the differential equations and equations of state which must be known in order to have achieved a basic understanding

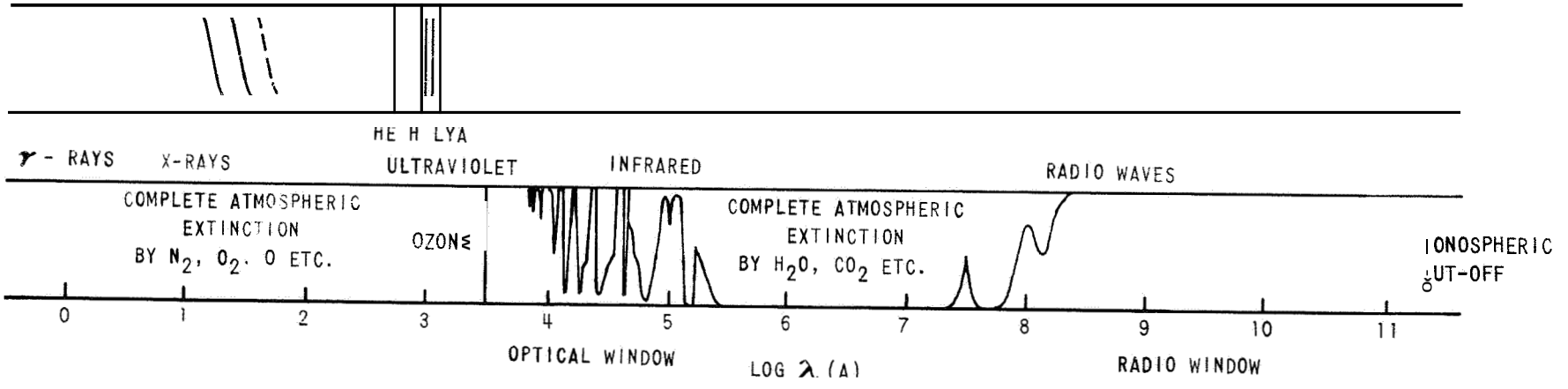


Figure 2 (S4) THE ELECTROMAGNETIC SPECTRUM

THE ELECTROMAGNETIC SPECTRUM IS EXHIBITED ON A LOGARITHMIC SCALE FROM X-RAYS TO THE RADIO SPECTRUM. THE ABSCISSA IS LOG A (ANGSTROMS). IN THE LOWER STRIP THE POSITIONS OF THE ATMOSPHERIC WINDOWS ARE INDICATED. THE TOP STRIP DEPICTS THE RANGE ACCESSIBLE TO DETECTORS FLOWN IN A SATELLITE ABOVE THE EARTH'S ATMOSPHERE. THE RADIATION FROM A DISTANT STAR IS ABSORBED BY THE HYDROGEN LYMAN LINES AND CONTINUUM AND BY THE HELIUM ABSORPTION PRODUCED BY THE INTERSTELLAR MEDIUM. NO SUCH BLOCKING IS PRODUCED FOR RADIATION FROM THE SUN. (REF. 1)

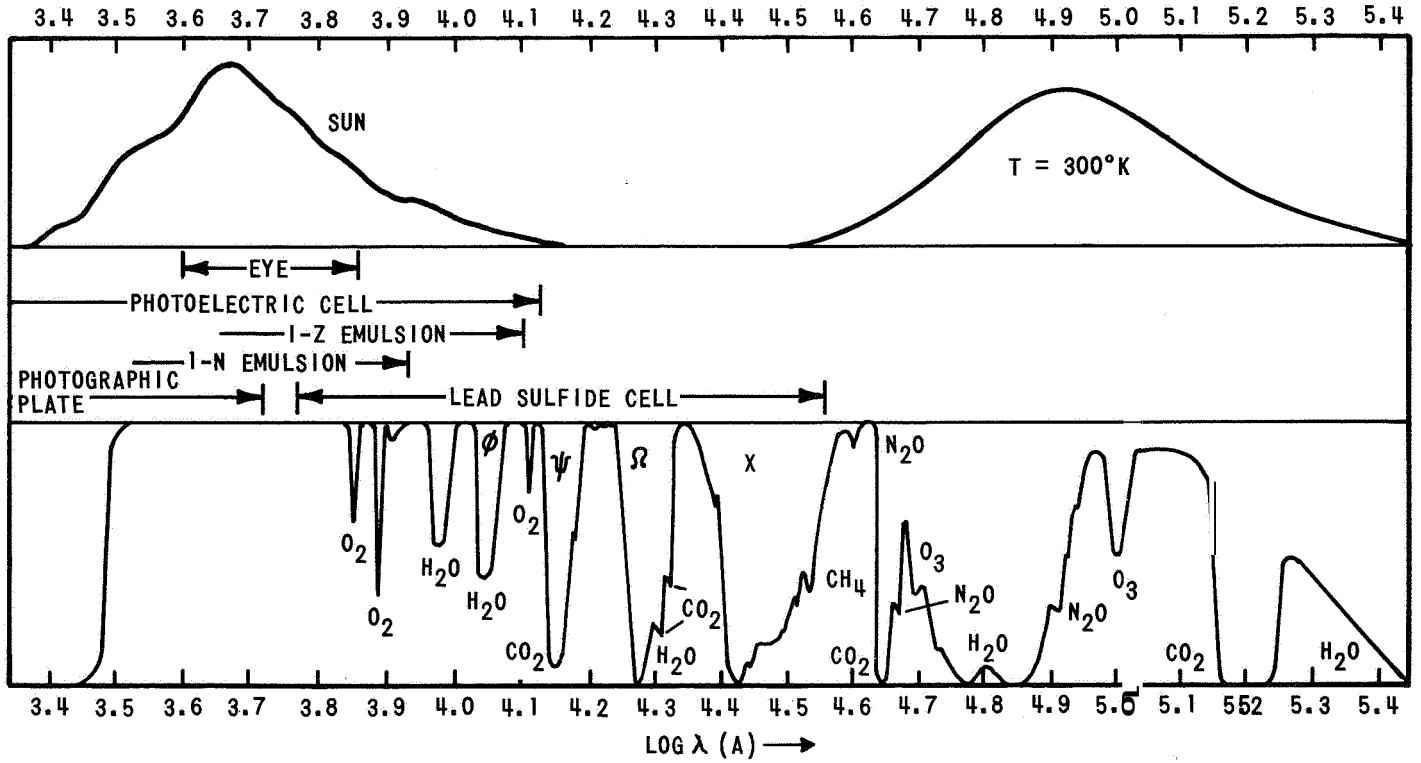


Figure 3 (S4) THE OPTICAL WINDOW IN THE EARTH'S ATMOSPHERE

THE REGION FROM 0.2 μ TO 25 μ IS SHOWN IN GREATER DETAIL. THE INFRARED-BAND ABSORPTION DUE TO WATER VAPOR AND OTHER ATMOSPHERIC CONSTITUENTS, HERE DEPICTED AS CONTINUOUS, ACTUALLY CONSISTS OF DISCRETE MOLECULAR LINES. IN SOME SPECTRAL REGIONS, THESE LINES OVERLAP SO BADLY AS TO PRODUCE A COMPLETE EXTINCTION OF THE OUTGOING RADIATION. THE APPROXIMATE SPECTRAL RANGES OF CERTAIN RADIATION DETECTORS IS EXHIBITED. THERMOCOUPLES, BOLOMETERS, AND GOLAY DETECTORS ARE SENSITIVE THROUGHOUT THE ENTIRE REGION.

THE TOP STRIP SHOWS THE RELATIVE ENERGY DISTRIBUTIONS IN THE SOLAR SPECTRUM AND IN A BLACK BODY AT 300°K (WHICH IS IN THE RANGE OF PLANETARY TEMPERATURES). (REF. 1)

of the physical picture of the atmosphere of a star in equilibrium. Since ideal assumptions are used and inherent errors exist in empirical data collection, discrepancies occur between the theoretical and experimental models. An applicable example is one where photospheric granulation suggests convective currents existing in the upper layers of the sun whereas the temperature gradient demanded by limb-darkening observations is more nearly in accord with that appropriate to the transport of energy by radiation[3]. The empirical paradox is unexplained.

1. Hydrostatic Equilibrium

The balance between the weights of various atmospheric layers and the pressures which support them defines hydrostatic equilibrium. Two competing forces, gravity and pressure, must balance since gravity would collapse the star if gas and radiation pressure did not suffice to keep it distended[2]. The mathematical representation expressing hydrostatic equilibrium is

$$\frac{dP}{dr} = -\frac{GM(r)}{r^2}\rho \qquad \text{(S4-1)}$$

For a unit volume, the pressure gradient, dP/dr, is equal but opposite the gravitational force $(GM(r)/r^2)\rho$, where G is the gravitational constant, 6.670×10^{-8} dynes/cm^2/g^2, ρ is the density and $M(r)$ is the mass interior to a sphere of radius r.

The mass gradient, $\nabla M(r)$ is readily derived by considering the mass of a spherical shell of radius r and thickness dr and is represented as:

$$dM(r) = 4\pi r^2 \rho\, dr \qquad \text{(S4-2)}$$

Using (S4-1) and (S4-2) we may combine the mass and pressure gradients in the form

$$\frac{1}{r^2}\frac{d}{dr}\left(\frac{r^2}{\rho}\frac{dP}{dr}\right) = -4\pi G\rho \qquad \text{(S4-3)}$$

an equation important in the study of convective zones and isothermal cores.

In equilibrium, the physical picture demands **two** additional gradients, $\nabla L(r)$ and ∇T, where $L(r)$ is the energy crossing a sphere of radius r per second and T is the temperature.

The luminosity gradient, $\nabla L(r)$, is an expression of the energy produced in a shell of mass $4\pi r^2 \rho\, dr$ and is

$$dL(r) = 4\pi r^2 \rho \varepsilon\, dr \qquad \text{(54-4)}$$

where ε is the energy produced per gram second.

The temperature gradient is directly responsible for energy transport and hence requires more exacting examination. There exist three transportation modes available for energy propagation: radiation, conduction, **or** convection. **For** the sun, conduction is negligible. Radiative energy transport will occur whenever there is a temperature gradient. The gradient must be large enough to drive the energy produced in the interior outward against the resistance set up by the opacity of the material. The higher the degree of attenuation introduced by the material, the steeper the temperature gradient must be. If the temperature gradient reaches **or** exceeds a specified critical value, radiation transfer decreases and convection becomes the dominant energy transfer mechanism; cf., equation (54-9).

2. Convection

The temperature gradient for radiative transport of energy

$$\frac{dT}{dr} = -\frac{3k_\nu \rho}{4ac}\frac{1}{T^3}\frac{L(r)}{4\pi r^2} \qquad \text{(S4-5)}$$

and for convective transport is

$$\frac{1}{T}\frac{dT}{dr} = \frac{\gamma-1}{\gamma}\frac{1}{P}\frac{dP}{dr} \qquad \text{(S4-6)}$$

For a rigorous derivation of the temperature gradients the reader is referred to Reference 2. In (S4-5) and (S4-6), a is the radiation constant equal to 7.568×10^{-16} ergs $cm^{-3} deg^{-4}$; c is the velocity of light, k_ν is the absorption coefficient (Rosseland mean opacity) in cm^2/gm; and γ is the ratio of specific heats, c_p/c_v which is taken as 5/3 for a perfect monatomic gas.

The "adiabatic exponent", γ , may be stated as

$$n = \frac{1}{\gamma-1} \qquad \text{(S4-7)}$$

where n is called the "polytropic index", then (S4-6) may be put into the form

$$\frac{1}{T}dT = \left(\frac{1}{n+1}\right)\frac{1}{P}dP \qquad \text{(S4-6a)}$$

or

$$\frac{d\log P}{d\log T} - 1 = n \qquad \text{(S4-6b)}$$

For an ideal monoatomic gas, $\gamma = 5/3$, $n = 3/2$, then

$$\frac{d\log T}{dr} = 0.4\frac{d\log P}{dr} \qquad \text{(S4-6c)}$$

K. Schwarzschild has established a criterion which states that the radiative temperature gradient is unstable against convection when it is steeper than the local adiabatic gradient. When pressure equilibrium is established in a small element due to a temperature, the matter expands,

creating a low density volume, which experiences a buoyant force driving the element outward. Stability occurs when the element rises to a region that has both pressure and temperature equilibrium. When the radiative gradient is steeper than the adiabat, the mass element departs from local temperature as it rises, and is continually driven outward. Thus when the radiative gradient is unstable, one must consider energy transport by convection as well as radiation. The total flux, set up by the temperature gradient, will be carried partly by radiation and partly by convection. The radiation flux will be

$$F_r = -\frac{4ac}{3k_\nu \rho} T^3 \frac{dT}{dr} \tag{S4-8a}$$

and the convective flux is

$$F_c = C_p \rho v \ell \left[\left(\frac{dT}{dr}\right)_{ad} - \left(\frac{dT}{dr}\right)\right] \tag{S4-8b}$$

where C_p is the specific heat at constant pressure, v is the velocity of the convective mass and ℓ is the distance a convective mass will travel before losing its identity and mixing with its surroundings. For the monatomic gas example of (S4-6c), we note that (S4-6b) is

$$\frac{d \log P}{d \log T} - 1 = 1.5 \tag{S4-6b}$$

Here the temperature gradient is set equal to the adiabatic gradient. The instability criterion is

$$\left|\left(\frac{dT}{dr}\right)_{rAD}\right| > \left|\left(\frac{dT}{dr}\right)_{AD}\right| \tag{S4-9}$$

or

$$n < 1.5$$

Since we have assumed an ideal gas (S4-6b) when integrated yields

$$PT^{-(n+1)} = \text{CONSTANT} \tag{S4-10}$$

which may be written as

$$P = K\rho^{\frac{n+1}{n}} = K\rho^{\gamma} \tag{S4-11}$$

For n constant (S4-11) is called a polytropic equation of state.

V. RADIATIVE EQUILIBRIUM

Radiative equilibrium implies that there are no mechanisms other than radiation for transporting heat in the atmosphere. Additionally, it is assumed that there are no sources of heat in the atmosphere. The effective temperature, T_e, electron pressure, P_e, local thermodynamic equilibrium (LTE) and the variation of temperature as a function of depth are the basic parameters which must be calculated in assuming radiative equilibrium.

Before examining the above relationships, a concept called Optical Depth and its correlation to geometric depth shall be reviewed.

1. Optical Depth

Consider a plane-layer atmosphere of homogeneous gases of thickness x. The initial radiation, I_ν^o, emitted perpendicular to the atmosphere becomes partially attenuated while transversing the thickness x of the atmosphere. In a distance dx, the intensity of the beam will be reduced by

$$dI_\nu = -k_\nu I_\nu \rho dx \tag{S4-12}$$

where k_ν is defined as the mass absorption coefficient for radiation of frequency ν. Integrating (S4-12) gives

$$\log I_\nu = -k_\nu \rho x + \text{CONSTANT}$$

$$I_\nu = I_\nu^o e^{-k_\nu \rho x} \qquad \text{(S4-13)}$$

Figure 4(S4) graphically illustrates the exponential reduction as the atmosphere increases in thickness.

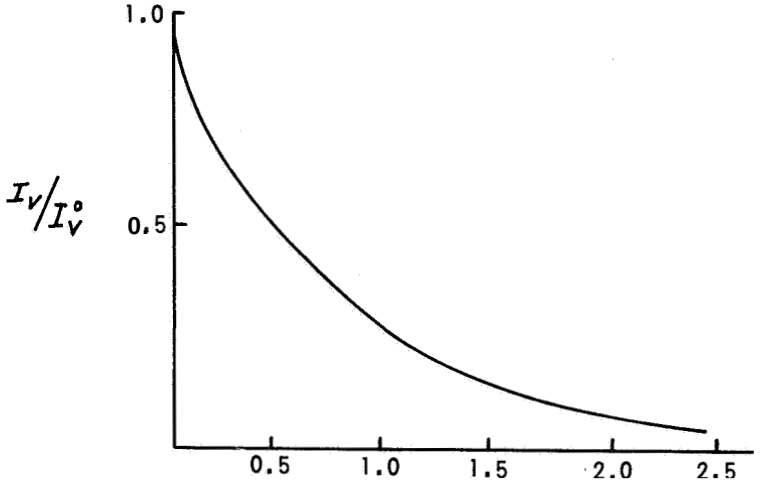


Figure 4 (S4) ATTENUATION OF RADIATION WITH DEPTH

The quantity τ is defined as the optical depth and is dependent on the absorption coefficient, k_ν as:

$$\tau = \int_o^x k_\nu \rho dx \qquad \text{(S4-14)}$$

As shown in Table I(S4) the larger the value of τ the greater the extinction in the atmosphere.

TABLE I(S4)

ENERGY ATTENUATION

τ	I_ν / I_ν^o
0. 1	0. 905
0. 5	0. 606
1. 0	0. 368
2. 0	0. 135
3. 0	0. 050
5. 0	0. 007

In the table, a unit optical depth corresponds to the distance in which the traversing radiation is reduced by 1/e. This value of $\tau_{5000} = 1$ defines the upper boundary of the photosphere at the limb, when viewed at 5000Å (Ref. 1). The correspondence between optical depth and geometric depth in kilometers [Fig. 5(S4)] exists assuming the solar gases in hydrostatic equilibrium corresponding to a boundary temperature T_o = 4400°K. The density is assumed as

$$\rho = \rho_o e^{-\mu g h / R_m T_o} = \rho_o e^{-h/90} \tag{S4-15}$$

where h is measured in kilometers, μ is the molecular weight (1. 50), R_m is the gas constant (8. 31 x 10^7 erg/(g mole)°K) and g is the gravitational constant. Over this small region, $k_\nu \approx \rho^{0.66}$

$$\tau = \int k_\nu \rho \, dh = \int \rho^{1.66} dh \sim e^{-\frac{1.66}{90} h} = e^{-h/54}$$

then

$$\tau = 2 \int_o^\infty k_\nu \rho \, dx = c e^{-h/54} \tag{S4-16}$$

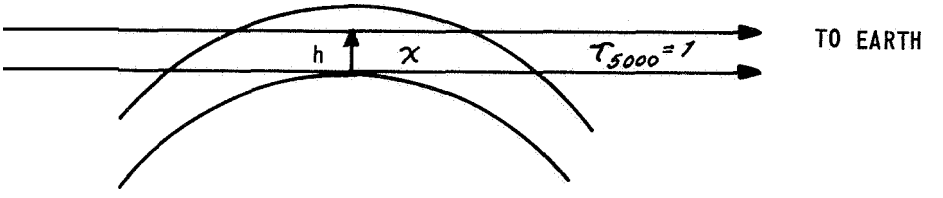


Figure 5 (S4) GEOMETRIC DEPTH (REF. 15)

letting $C = 1$, requires $h = 0$ for $\tau = 1$. The observed intensity is now:

$$I = I_o \int_{\tau}^{o} e^{-\tau} d\tau = I_o (1 - e^{-\tau}) \qquad \text{(S4-17)}$$

as shown in Table II(S4).

The photosphere appears to be transparent to -380 km below the **zero** level. As the line of sight moves from the solar limb to the center of the disk, $\tau_{5000} = 1$ corresponds to a height of -300 km. This would imply an equivalent optical depth of $\tau_{5000} = 0.003$ above the zero level when the sun is viewed radially. Thus while the limiting temperature for $\tau_{5000} \cong 0$ is 4300°K, the temperature along the line of sight at the limb increases up to 4600°K.

TABLE II(S4)

OPTICAL AND GEOMETRIC DEPTH
AT THE SOLAR LIMB

h	τ_{5000}	$I_{\lambda}/I_{\lambda}^{\circ}$
162 km	0. 05	0. 05
108	0. 14	0. 13
54	0. 37	0. 31
0	1. 00	0. 63
-54	2. 72	0. 93

2. Effective Temperature

The assumptions defining radiative equilibrium retain their validity by demanding that the net flux of radiation, πF, summed over all wavelengths, be constant. The emergent flux will have the same value πF. [4] We now introduce an effective temperature T_{ε} related to the constant net flux πF by

$$\sigma T_{\varepsilon}^{4} = \pi F \qquad \text{(S4-18)}$$

where σ = 5. 75 x 10^{-5} erg sec $^{-1}$cm^{-2}K^{-4}, is the Stefan-Boltzmann constant, and $\sigma T_{\varepsilon}^{4}$ is the Stefan-Boltzmann law for a black body. The flux of the emergent radiation depends on the rate at which the temperature increases with optical depth. The effective temperature, T_{ε}, is the temperature of a black body that would radiate the same total amount of energy that a particular body does, and for this reason is the physically significant temperature compared to the color temperature, T_{c}, which is the temperature of a star as estimated from the intensity of the stellar radiation at two or more wavelenghts. T_{c} amounts to fitting a Planckian curve to the flux, over a definite interval, 0. 4 μ - 0. 65 μ, and utilizing this curve as a convenient interpolation device. The relationship of T_{c} to T_{ε} will depend on the relationship of T_{c} to both the temperature gradient and the absorption coefficient.

3. Electron Pressure

The electron pressure, P_ε, is the rate at which ions can re-capture electrons and would be more suitably called the electron density. It is expressed as

$$P_\varepsilon = k N_\varepsilon T \tag{S4-19}$$

where N_ε is the number of electrons per unit volume, k is Boltzmann's constant = 1.38×10^{-16} ergs K^{-1} and T is the temperature in degrees Kelvin.

Equation (S4-19) is the equation of state for a perfect gas written in an alternate form other than

$$PV = nRT, \qquad P = \rho\left(\frac{k}{\mu H}\right)T \tag{S4-20}$$

where μ is the total molecular mass, and H is the mass of *a* particle of unit atomic weight.

For a given temperature and density, in addition to knowing the relative number of atoms in various excited levels, information pertinent to the relative number of neutral and ionized atoms is also desired. Saha, derived an ionization formula which when employed in determining stellar atmospheres, takes on the form[5]

$$\frac{N_{q+1}\, N_\varepsilon}{N_\varepsilon} = \frac{(2\pi m k T)^{3/2}}{h^3} \frac{2\mathcal{U}_{q+1}(T)}{\mathcal{U}_q(T)} e^{-\chi_q/kT} \tag{S4-21}$$

where N_q is the number of atoms in the q^{th} state, N_{q+1} is the atoms in the $q+1$ state, χ_q is the ionization energy, in electron volts, necessary to ionize the atom from the q^{th} to the $(q+1)$-st state and $\mathcal{U}_q(T)$ are the partition (probability) functions for the ionized state.

Substituting (S4-19) and (S4-21) and letting q = the ground state, gives

$$\frac{N_1 P_e}{N_0} = \frac{(2\pi m)^{3/2}}{h^3}(kT)^{5/2}\,\frac{2\mathcal{U}_1(T)}{\mathcal{U}_0(T)}\,e^{-\chi_0/kT} \qquad \text{(S4-22)}$$

For numerical calculations (S4-22) is put in the logarithmic form

$$\log_{10}\left(\frac{N_1 P_e}{N_0}\right) = -\frac{5040}{T}\chi_{0,1} + 2.5\log_{10}T - 0.48 + \log_{10}\frac{2\mathcal{U}_1(T)}{\mathcal{U}_0(T)} \qquad \text{(S4-23)}$$

where P_e is the electron pressure in dynes cm^{-2}, N_1 is the number of ionized atoms cm^{-3}, N_0 is the number of neutral atoms cm^{-3} and

$$\log_{10}\left\{\frac{(2\pi m)^{3/2}}{h^3}k^{5/2}\right\} = -0.48\,, \quad \log_{10}e^{\frac{-\chi_{0,1}}{kT}} = -\frac{5040}{T}\chi_{0,1} \qquad \text{(S4-24)}$$

The ionization theory's outstanding achievement was its interpretation and prediction of the spectral lines emitted at a particular temperature and density.

4. Local Thermodynamic Equilibrium (LTE)

For LTE to exist, it is assumed that one can define, at each point in the atmosphere, a local temperature T. This local temperature is the electron temperature corresponding to Planck's function, $B_\nu(T)$

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\,\frac{1}{e^{h\nu/kT}-1} \qquad \text{(S4-25)}$$

where h is Planck's constant and ν is the frequency.

After an atom absorbs a photon of frequency ν, and before it can re-emit, it may have a collision with a free electron or hydrogen atom. Then the absorbed energy is converted into kinetic energy and eventually re-emitted at a different frequency. The interaction between the free electrons

and the excited states of the atom, insures a Boltzmann population of atomic levels, B_ν, and additionally leads to the fact that the re-emission, J_ν, is governed by Kirchoff's law

$$J_\nu = k_\nu B_\nu(T) \tag{54-26}$$

where j_ν is the emission coefficient, k_ν is the absorption coefficient (Section VI-1) and $B_\nu(T)$ is Equation (S4-25) defined as the source function. The LTE assumes pure absorption and the non-existence of scattering in the atmosphere. The mechanisms which generate spectral lines have been found to be a combination of pure absorption and coherent scattering, both of which shall be discussed in greater detail in Section VI.

5. Limb Darkening

The intensity of radiation at the Sun's limb is less than the intensity measured at the center of the solar disk, and hence the limb appears darkened. The darkening of the Sun's limb is a consequence of the temperature gradient in the outer layers of the atmosphere. Minnaert[6] has stated that the brightness decrease toward the limb and the associated color reddening implies that the decrease in brightness is more pronounced for the shorter wavelengths, that the solar layers are transparent and that the temperature increases monotonically with depth. "In the center of the disk the observer perceives radiation from deeper layers, near the limb from more superficial layers. Quantitative study of the limb darkening is the classical method by which vertical temperature distribution in the photosphere is derived. However, very high precision is required for the derivation of a detailed model; and the temperature in the higher layers can be found only from measurements extended to the extreme limb, which are difficult to make."

The observed brightness, denoted by $I_\lambda(\tau, \theta)$ where τ is given in Equation (54-14) and θ is the angle between the normal to the surface and the line of sight, is given by the summation of the source functions for the element of overlying layers, i. e.,

$$I_\lambda(0,\theta) = \int_0^\infty B_\lambda(T_{\tau_\lambda})\, e^{-\frac{\tau_\lambda}{\cos\theta}}\, \frac{d\tau_\lambda}{\cos\theta} \qquad \text{(S4-27)}$$

where τ_λ is defined in Equation (S4-14). The ordinary procedure in obtaining a solution of (S4-27) is to set

$$B_\lambda = a + b\tau_\lambda + c\tau_\lambda^2 + \cdots \qquad \text{(S4-28)}$$

and by direct integration obtain

$$I_\lambda = a + b\cos\theta + 2c^2\cos^2\theta + \cdots \qquad \text{(S4-29)}$$

The coefficients may be obtained from direct observation. Table III(S4) obtained from Reference 6, depicts an empirical model of temperature variations as a function of optical depth, as derived from the observed radiance.

For $\lambda = 5010\text{Å}$, $\cos\theta \sim \tau_{5000}$

$$I_{5010} = (4.151 \times 10^{14})(0.259 + 0.872\tau - 0.133\tau^2) \qquad \text{(S4-30)}$$

and for $\lambda = 10,080\text{Å}$, $\cos\theta \sim \tau_{10,080}$

$$I_{10,080} = (1.151 \times 10^{14})(0.563 + 0.600\tau - 0.163\tau^2)$$

The model is obtained by generating the value of I_λ in (S4-30) or B_λ in (S4-28) and solving for T in (S4-25) for each I_λ or B_λ .

TABLE III(S4)

EMPIRICAL MODEL (REFERENCE 6)

τ_{5010}	I_{5010}*	Temperature	$\tau_{10\,080}$	$I_{10\,080}$
0. 28	2. 04 x 10^{14}	5500°K	0.42	0. 86 x 10^{14}
0. 62	3. 16	6000	0. 85	1. 07
1. 07	4. 57	6500	1.40	1. 28
1. 68	6. 18	7000	2. 19	1.49

* Units: erg · cm^{-2} · sec^{-1} · $ster^{-1}$ per unit wavelength.

6. Temperature and Depth

For an atmosphere whose attenuation function is not dependent on the wavelength utilized in the observation, i. e., a gray atmosphere, the temperature may be approximated by the empirical relationship

$$T^4 = \frac{1}{2} T_\varepsilon^4 \left(1 + \frac{3}{2}\tau\right) \qquad \text{(S4-31)}$$

where T_ε is the Sun's effective temperature (5760°K) and τ is defined by (S4-14).

As stated in Section V-1 the upper boundary of the photosphere is defined as that level at which the optical depth, viewed at 5000Å along the line of sight at the extreme limb is unity. A unit optical depth corresponds to the distance in which the traversing radiation is attenuated by a factor 1/e. The figure below shows a relationship between the optical depth, temperature, and height or thickness of the photospheric layer at the limb.

For an optical depth, τ = 1, the corresponding temperature is 6000°K, and the height is 0 kilometers, thus defining the upper boundary of the photosphere. Values of $\tau > 1$ refer to increasing temperature regions. For values of $\tau < 1$, the corresponding heights up to +360 km (τ = 0), the temperature distribution immediately defines the reversing layer.

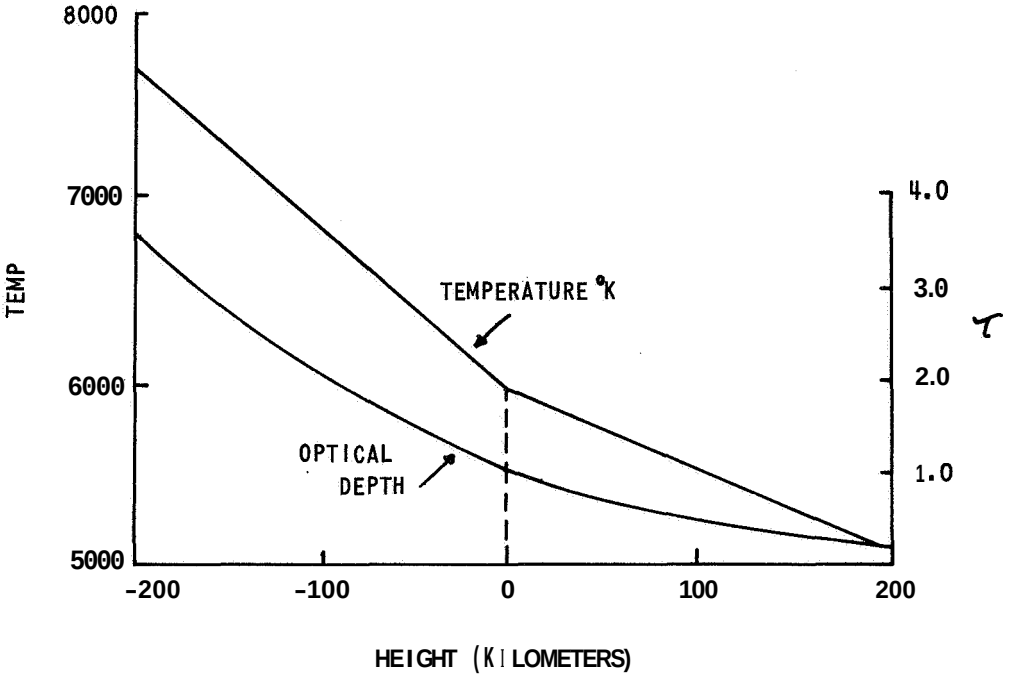


Figure 6 (S4) OPTICAL DEPTH, TEMPERATURE AND HEIGHT (REF. 11)

The reversing layer, is thought to explain two solar phenomena: the dark Fraunhofer lines and the bright lines known as the "flash spectrum". The Fraunhofer lines are dark absorption lines caused by photons of light traversing a cooler gas. Those wavelengths which are characteristic of the cooler gas are selectively absorbed out of the traversing radiation by atoms and molecules situated in the same layers in which the photospheric spectrum is formed, thus producing a dark line spectrum. During a solar eclipse, the spectrum is reversed thus giving a "flash spectrum". The mechanism which integrates and controls these functions has been found to be located in the extreme upper band of the photosphere (reversing layer). The darkening of the limb is attributed to a decrease in temperature. This concept becomes more pronounced, when one examines the correlation existing between the height (depth) of the isothermal layers and their associated temperatures.

TABLE IV(S4)

HEIGHT AND TEMPERATURE CORRELATION

Height (km)	-300	-100	0	500	1000
Temperature (°K)	6000	4700	4300	6000	7000

The formulation which delineates the sharp demarcation line as defined by $\tau = 1$, generates sequential depth values. DeJager (1952), Pierce, Aller (1951) have made careful temperature determination at one wavelength ($\lambda = 5000\text{Å}$), as a function of optical depth [Table V(S4)].

TABLE V(S4)

TEMPERATURE AS A FUNCTION OF OPTICAL DEPTH

τ_o	$e^{-\tau}$	DeJager (1952)°K	Pierce, Aller (1951)°K	Ten Bruggencate (1950)°K
0.003	.997	-	-	-
0.01	.99	4790	4600	4620
0.02	.98	4820	4700	4790
0.04	.96	4890	4810	4980
0.06	.94	4960	4900	5110
0.08	.92	5040	5000	5230
0.10	.905	5100	5100	5320
0.20	.819	5350	5400	5380
0.40	.67	5750	5770	5620
0.60	.549	6010	6060	-
1.0	.368	6410	6460	-
1.4	,247	6730	6760	-
1.8	.165	6990	7050	-

Except in the outermost layers, the agreement is excellent. The discrepancies in the temperature layers are liable to be from two sources in the empirical model: (1) the uncertainty of the measurements of the radiance near the limb and (2) granulation may produce a billowy appearance of the solar surface [Figure 7(S4)].

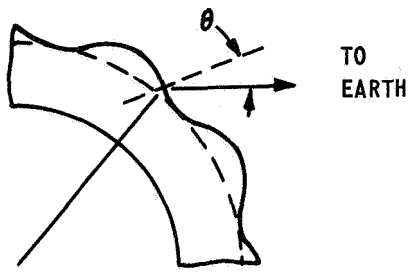


Figure 7 (S4) BILLOWY GRANULATION (REF. 6)

VI. ELEMENTAL ABUNDANCES

The third assumption, and perhaps the most critical, is the knowledge of the elemental abundances and the consequences of that knowledge, i. e., the hydrogen-to-metal ratio. Subdivision VI will cover the theoretical determination of the absorption coefficients, f-values, and damping constants. Empirically determined parameters will comprise the final half of this section. Equivalent widths will be discussed including curves of growth plus the line formation which is indicative of line intensity, including radiation transfer, line absorption, coherent scattering and blanketing. The discussion of the physical state of line intensity including temperature, pressure, turbidity, chemical composition and electric and magnetic fields will be followed by defining the necessary atomic parameters governing line intensity, such as transition probability, collision area, Zeeman g-factor and hyperfine structure. Sections on isotope abundance and molecular dissociation potential shall terminate this subdivision.

1. Absorption Coefficient

The hydrogen-to-metal ratio in the Sun is found from the intensities of metallic lines produced in an atmosphere whose continuous absorption is due primarily to the negative hydrogen ion[8], in the wavelength interval 0.4μ - 2.0μ. The solar atmosphere is composed of one metal atom to about 10^4 hydrogen atoms. For $T \approx 6000°K$, each metal atom releases one electron which may attach itself to one of the abundant hydrogen atoms to form a negative hydrogen ion, H^-. Thus the concentration of H^- is proportional to the electron pressure, P_e. The outer electron has only one stable orbit from which it may be removed by the absorption of radiation, the ionization potential, $\chi_{g,g+1}$, being 0.754 ev.

Saha's equation (S4-23) indicates that for a temperature of $T \approx 6000°K$, the ratio of the number of negative hydrogen ions to hydrogen atoms is $N_{H^-}/N_H = 10^{-7}$. The absorption spectrum of H^- in its ground state is a continuum whose edge falls at the wavelength corresponding to 0.754 ev. (1.645μ), rises steadily to a maximum near 0.85μ, and decreases at shorter wavelengths. This is the "bound-free" absorption, so named because the principal contribution to the continuous absorption is made by photo-ionization from the third bound level to the continuum. From Figure 8(S4), it can be seen that the continuous absorption increases after reaching an extremum at 1.645μ. The increase is caused by the "free-free" absorption, which is due to energy exchange of unquantized electrons at high temperatures and thus becomes important in the infrared.

That the continuous absorption is temperature dependent becomes clearly established in Figure 8(S4). It illustrates the variation of the continuous absorption coefficient with wavelength in the solar atmosphere. Here k_γ is the absorption coefficient and $\bar{k}$ is a mean absorption coefficient. The analysis of the continuous spectrum of the sun shows that the coefficient of continuous absorption in the solar atmosphere increases by a factor of the order of 2 as the wavelength increases from 4000Å to 9000Å; beyond 9000Å, the absorption coefficient decreases until at about 16,000Å, it has a pronounced minimum; and beyond 16,000Å the absorption coefficient increases again (Ref. 4).

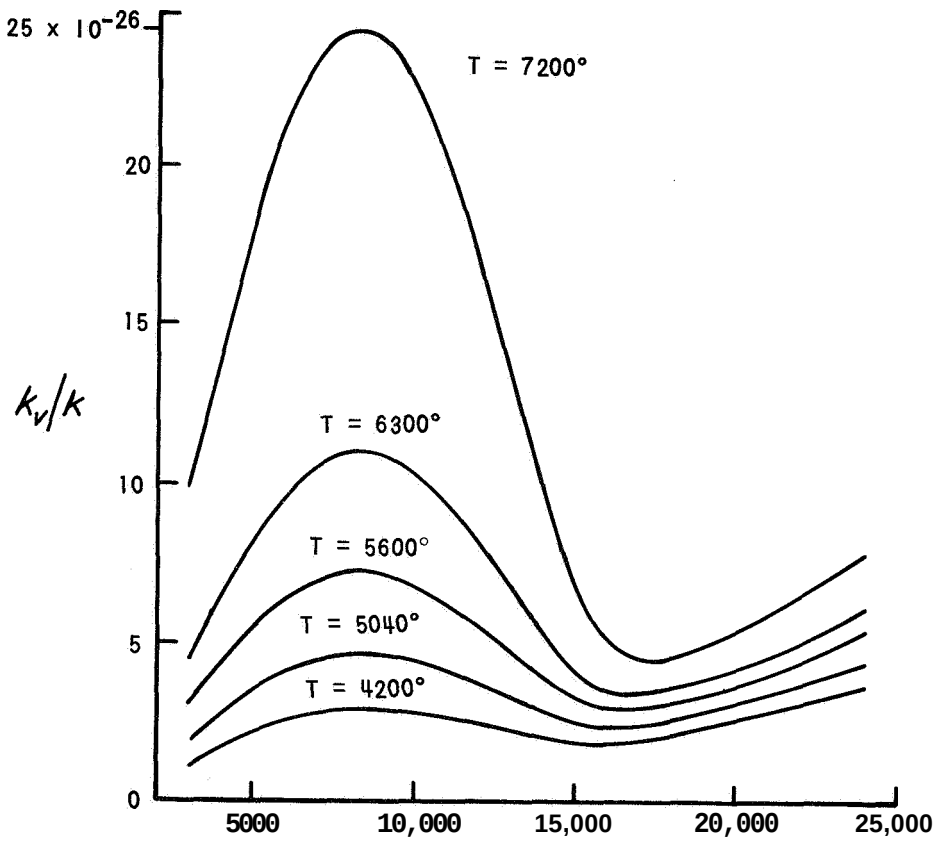


Figure 8 (S4) THE CONTINUOUS ABSORPTION OF H^-. THE ABSCISSA DENOTE THE WAVE-LENGTH IN ANGSTROMS AND THE ORDINATES REPRESENT THE ABSORPTION COEFFICIENT PER NEGATIVE HYDROGEN ION AND PER UNIT ELECTRON PRESSURE AT THE VARIOUS TEMPERATURES INDICATED, (REF. *6*).

2. Oscillator Strengths (f-values)

The Ladenburg F or f-value is the number of electrons oscillating at a frequency ν_o per atom. Only for the transitions in hydrogen is it possible to calculate the f-values exactly since the knowledge of a satisfactory wave function is of primary importance. The first few members of the Balmer series have f-values as follows:[9]

$H_\alpha = 0.6408, \quad H_\beta = 0.1193, \quad H_\gamma = 0.0447,$
$H_\delta = 0.0221$

The f-values or oscillator strength $f_{n,n'}$ are related to the Einstein spontaneous emission coefficient, $A_{n,n'}$, by

$$A_{n,n'} = \frac{g_{n'}}{g_n} \frac{8\pi^2 e^2}{m^2 \lambda^2} f_{n,n'} \qquad \text{(S4-32)}$$

or

$$f_{n,n'} = 1.5 \times 10^{-8} \lambda^2 \frac{g_n}{g_{n'}} A_{n,n'}$$

where λ is expressed in microns. Here g_n and $g_{n'}$ are the statistical weights of the upper and lower population levels, respectively. The knowledge of f-values is necessary for the interpretation of line intensities. In complex atoms, such as iron, accurate f-values cannot be calculated by theoretical means, and recourse must be had to empirical determinations of line strengths and f-values. Determination of relative f-values requires a knowledge of the temperature of the absorbing source of line spectra, but the total number of participating atoms in the line of sight need not be known.

A measurement of the equivalent width, W, for a particular absorption line, enables one to determine the relative f-values, Nf from the relation

$$W \propto \lambda_o^2 (Nf) \qquad \text{(S4-33)}$$

where N is the particle density. f-values for elements have also been

determined in laboratory experimentation using electric furnaces and atomic beam methods[12] with the results being applied to the solar spectrum.

3. Damping Constants

The broadening of a spectral line depends in addition to natural and Doppler broadening, on the radiation damping constant, which may be calculated from the f-values and radiation field, and the collisional damping constant which is governed by the temperature, density and character of the perturbing ions or atom. Damping constants are determined experimentally, in the laboratory, and applied to the solar spectrum.

The radiative damping constant, $\gamma_r = 0.2223/\lambda^2$ (sec^{-1}) is the reciprocal of the mean of the radiative life times of two energy states, in sec^{-1}. The values of γ_r are obtained from tables of transition probabilities. For most absorption lines γ_r is of the order of 10^8. Observations have indicated that the collisional damping constant, $\gamma_c = 2/\tau$, where τ is the "mean free time" between collisions, exceeds the radiation damping constant by factors of 5 to 100 in the moderate and deep layers of the photosphere, but in the higher layers radiative damping must be included and γ_r is replaced by $\gamma_r + \gamma_c$. "Although the theory of collisional damping is not sufficiently refined for the accurate calculations of γ_c, it can be combined with laboratory experiments to give reasonably useful results."[13]

4. Equivalent Width

Spectral lines are not perfectly sharp, but have a finite width due to various broadening mechanisms. The profile is the shape of the line as seen by a microphotometer measuring directly the intensity of the light in the spectrum at various wavelengths.[10] Because the profiles of different lines differ in shape, it is convenient to define some measurable quantity that can be used to calculate the total amount of light energy subtracted from the continuum by absorption. The measurable quantity utilized is the equivalent width, W, which is the width of a hypothetical line with rectangular profile of zero intensity along its entire width [see Fig. 9(S4)]. The equivalent width, W, represents the same subtraction of light from the spectrum as is removed by the actual line.

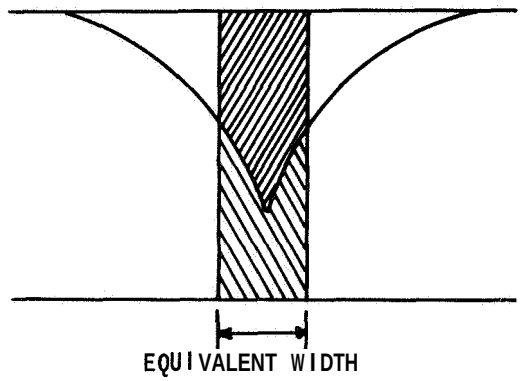


Figure 9 (S4) EQUIVALENT WIDTH OF A LINE. THE AREA OF THE RECTANGLE IS EQUAL TO THE AREA UNDER THE LINE PROFILE (REF 10).

Converting line profiles to equivalent widths has the advantage that they are relatively easy to measure, are not affected by the instrumental profile, and are expressed as single numbers giving the amount of light abstracted from the continuum. The disadvantage is that the equivalent width may hide some infor-mation which is necessary to interpret the structure of the atmosphere.

5. Curves of Growth

Since an absorption line arises from atoms at all depths throughout the photosphere, the losses to atoms at all layers combine to give the observed profile or equivalent width. It has been found that for a low elemental abundance, the equivalent width of a line is proportional to the product of the f-value and density of the atoms of the element present, Nf. As the elemental abundance is increased, the equivalent width varies as the square root of Nf. A graph of the equivalent widths plotted against the number of atoms that create the line profile is called a curve of growth. Figure 10(S4) shows the appearance of a typical curve of growth.

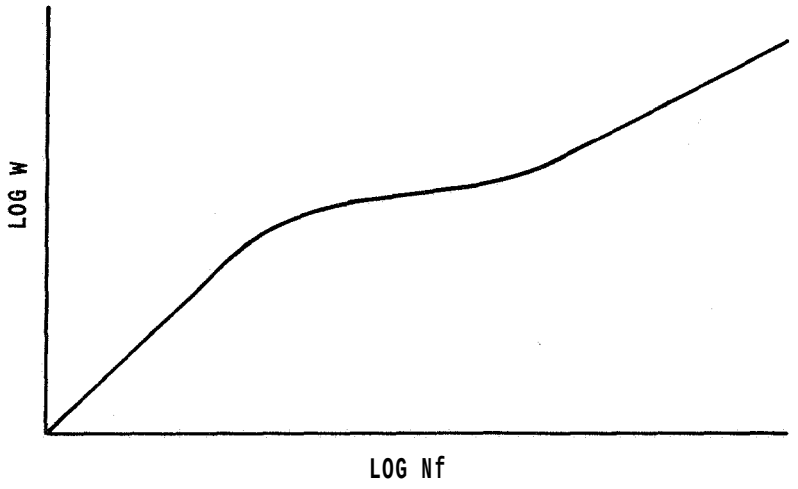


Figure 10 (S4) A CURVE OF GROWTH. LINES ON THE FLAT PORTION OF THE CURVE EXHIBIT A BELL-SHAPED PROFILE. THE LARGER THE DAMPING CONSTANT THE SMALLER THE FLAT PORTION WILL BE (REF 10).

6. Line Intensity

The intensity of the spectral line is dictated by the method by which the line is formed, the immediate physical state of the atmosphere and the element forming the line. Line formation, may be caused by pure absorption, pure scattering or both. There are certain lines for which one process or the other predominates. In addition to the continuous absorption examined in Section VI-1, there exists a line absorption coefficient per gram, k_ℓ The calculation of k_ℓ , usually involves a variety of uncertain assumptions. The line absorption coefficient is proportional to $n_{r,s}$ the number of atoms per gram in state of ionization r and level of excitation s , and is related to the total number of atoms per gram through the Saha and Boltzmann equations, and therefore depends on the temperature and electron pressure, which are functions of the depth τ_λ . It is also frequency

dependent and is therefore associated with the mechanisms of line broadening. The quantity n_ν is defined as the ratio, k_ℓ / k_ν, of the line to the continuous absorption coefficient.

The physical state, in addition to the influence of temperature, pressure and chemical composition, must also encompass turbidity, due to convective mass motion and hyperfine line splittings, caused by electric (Stark) and magnetic (Zeeman) fields. Such are the variables which must be known or assumed.

7. Isotope Abundance

The isotope abundances or isotope ratios of the elements do not directly enter into the solar atmospheric model. A knowledge of the presence of isotopes aids the builders in the labeling of unidentified faint spectral lines. The search for isotopes in the Sun is made by means of an analysis of molecular (band) spectra, and as yet has been unsuccessful[7]. It has been suggested that the isotope doublets of $C^{13}N^{14}$ are sufficiently well resolved from $C^{12}N^{14}$ as to be detectable. Similarly, He^3, D/H, and Li^6/Li^7 have been sought, all to no avail. Of the twenty-six thousand lines in the photographic region, about 70 percent have been either wholly or partially identified with the general consensus of opinion accepting the hypothesis that the remaining 30 percent will prove to be of molecular origin.

8. Dissociation Potential

Molecular bands may be used to determine the elemental abundances, provided the f-values and molecular dissociation potentials are known. The latter are difficult to obtain and are as yet unknown for many astrophysically important molecules such as CO, CN, N_2 and NO. Since the Sun is not a sufficiently cool star, molecular bands are weak and not considered as an abundance determining method.

VII. MODEL BUILDING

Many solar atmospheric models have been proposed where each one specifies the electron pressure and temperature as a function of the optical depth at a standard wavelength $\lambda_{5000°}$. According to the assumptions and governing equations encountered in Sections IV, V and VI, one would expect a model to be as follows: a semi-infinite gaseous layer with a plane boundary in which the temperature, pressure and absorption coefficients at each depth are determined by the chemical composition and net flux.

The calculation of a detailed photospheric model proceeds in successive steps as follows:

(1) The degree of ionization as a function of P_e and T will be calculated according to Saha's equation (S4-21). When the ionization is very small,

$$\log_{10} \frac{N_{q+1}}{N_q} \simeq \log_{10} N_{q+1} \quad and \quad N_{q+1} \propto \frac{1}{P_e} \tag{S4-33}$$

For ionization that is almost complete

$$\log_{10} \frac{N_{q+1}}{N_q} \simeq -\log_{10} N_q \quad and \quad N_q \propto P_e \tag{S4-34}$$

(2) The relationship between total gas pressure, P_g, and electron pressure, P_e, is de'pendent on the hydrogen-to-metal ratio, A. In the Sun the bulk of the gas pressure is supplied by hydrogen, whereas P_e comes solely from ionization of metals as derived from the assumed proportions of various elements. A most crucial parameter is the H/metal ratio presented in Table VI(S4). The H/metal ratio, A, is calculated as 1.35×10^4, or Log A = 4.13. The calculation of $P_g = P_g(P_e, T)$ is accomplished as follows: χ_E is denoted as the fraction of atoms of element E that have been ionized, N_a, the number of atoms, both neutral and ionized, of all types, and N_e the number of electrons/cm^3. Then

$$\frac{P_g}{P_e} = \frac{(N_a + N_\varepsilon)\, kT}{N_\varepsilon\, kT} \qquad (S4\text{-}35)$$

and the problem reduces to computing N_e/N_a. By selecting a value of P_e, we may calculate χ_ε for each element. For each χ_ε known, N_e/N_a can be found and the ratio P_g/P_e will then be determined.

TABLE VI(S4)

RELATIVE ELEMENTAL ABUNDANCES ADOPTED FOR THE SUN (REF. 1)

Log A = 4.13

Element	Abundance	Element	Abundance
H	1,000,000	Na	2. 0
He	140, 000	Mg	19. 0
C	250	Al	1. 6
N	410	Si	40. 0
O	1, 300	K	0. 1
Ne	2,000	Ca	2.4
		Fe	6. 0
		Ni	0. 6
		Misc	2. 0
			73. 7

Table VII(S4) is the final result. Pressure is in dynes/cm^2, $\theta = 5040/T$.

(3) Calculation of the absorption coefficient as a function of P_e and T.

(4) Calculation of the "mean" absorption coefficients, $\bar{k}$, as a function of P_e and T. T is taken as 5600°K or $\theta = 0.9$, $P_e = 13$ dynes/cm^2.

Table VII (S4)

RELATION BETWEEN GAS AND ELECTRON PRESSURE FOR A SOLAR-LIKE ABUNDANCE (REF. 1)

LOG P_e	LOG P_g							
	θ = 0.50	θ = 0.60	θ = 0.70	θ = 0.75	θ = 0.80	θ = 0.85	θ = 0.90	θ = 0.95
-2.0							- .61	.07
-1.5						- .22	+ .36	1.07
-1.0				- .46	- .02	+ .63	1.34	2.04
- .5	- .17	- .16	- .05	+ .30	+ .91	1.62	2.33	2.98
0	+ .33	+ .34	+ .65	1.19	1.88	2.60	3.28	3.84
+ .5	+ .83	.87	1.48	2.15	2.87	3.57	4.19	4.58
1.0	1.34	1.45	2.41	3.12	3.84	4.51	4.99	5.26
1.5	1.84	2.14	3.38	4.11	4.81	5.38	5.72	5.94
2.0	2.37	2.96	4.36	5.08	5.72	6.16	6.43	6.66
2.5	2.94	3.89	5.35	6.03	6.56			
3.0	3.62	4.86	6.32					
	LOG P_g							
	θ = 1.0	θ = 1.1	θ = 1.2	θ = 1.4	θ = 1.6	θ = 1.8	θ = 2.0	
-2.0	.78	1.92	2.22	2.65	3.27	3.64	4.20	
-1.5	1.75	2.60	2.78	3.36	3.89	4.33	5.06	
-1.0	2.65	3.20	3.39	4.08.	4.53	5.09	5.95	
-0.5	3.46	3.79	4.06	4.76	5.22	5.95	6.86	
0	4.14	4.42	4.76	5.41	5.97	6.85	7.85	
+0.5	4.78	5.10	5.50	6.09	6.82			
1.0	5.44	5.83	6.22	6.83				
1.5	6.14.	6.58						

(5) Integration of $K(P_e, T)$ and $P_g(P_e, T)$. The integrated functions are derived from the hydrostatic equilibrium equation (S4-1).

$$dP = g\rho dx \tag{54-36}$$

and the optical depth (τ), equation (S4-14)

$$d\tau = k_\nu \rho \, dx \tag{S4-37}$$

Then

$$\frac{dP}{d\tau} = \frac{g}{k_\nu} \tag{S4-38}$$

is the equation which must be satisfied to obtain the variation of gas pressure with τ .

Since K_ν depends on P_e and T , we can find $K(P_g, T)$ from $K(P_e, T)$ and $P_g(P_e, T)$ both of which depend in a known fashion on the assumed composition. "Therefore, at the outset we must postulate the chemical composition of the Sun, or at least A , the ratio of hydrogen to the metals by number of atoms. Also, it is necessary to specify B , the He/H ratio by the number of atoms. Helium does not contribute to the opacity but adds to the total weight of the material in the layers. "[14]

An initial approximation (all metals are assumed singly ionized) that is used in finding P_g is

$$P_g = P_E A(1+B) \tag{S4-39}$$

where B is the He/H ratio, 1/7.

(6) Calculate the density, which is derived from the ideal gas law, using

$$\rho = \frac{P}{RT} = \frac{P_\mu}{R_m T} \tag{S4-40}$$

where R_m = 8. 31 x 10^7 ergs/(g mole)°K.

As an example of a theoretical photospheric model, the results obtained by Munch (1947) have been selected and are shown in Table VIII(S4).

Table VIII (S4)

THEORETICAL MODEL BY MUNCH (1947) (REF. 15)

τ	LOG P	LOG P_e	T (°K)	θ=5040/T	ρ (Gm/Cm^3)	t (Km)	LOG $\frac{N_g+1}{N_g}$
0.01	3.74	9.85-10	4650	1.080	1.4×10^{-8}	515	3.89
0.02	4.01	0.09	4700	1.074	2.7	428	3.92
0.04	4.19	0.27	4740	1.062	3.9	370	3.92
0.06	4.30	0.36	4790	1.051	5.0	333	3.94
0.08	4.38	0.44	4840	1.041	6.0	307	3.94
0.10	4.43	0.50	4890	1.031	6.6	290	3.93
0.12	4.46	0.56	4930	1.022	7.0	280	3.90
0.14	4.51	0.61	4970	1.014	7.8	262	3.90
0.16	4.54	0.64	5010	1.006	8.3	252	3.90
0.18	4.57	0.68	5050	0.998	8.8	242	3.89
0.20	4.60	0.71	5090	0.991	9. 4	232	3.89
0.24	4.64	0.77	5160	0.977	10.1	218	3.87
0.28	4.68	0.82	5220	0.965	11.0	203	3.86
0.32	4.71	0.87	5290	0.953	11.7	192	3.84
0.36	4.74	0.92	5350	0.942	12.4	181	3.82
0.40	4.77	0.96	5400	0.933	13.1	170	3.81
0.50	4.82	1.06	5540	0.910	14.4	151	3.76
0.60	4.86	1.15	5660	0.891	15.4	135	3.71
0.70	4.89	1.24	5770	0.874	16.2	123	3.65
0.80	4.91	1.32	5870	0.858	16.6	115	3.59
0.90	4.93	1.40	5970	0.844	17.2	107	3.53
1.0	4.94	1.48	6070	0.831	17.3	103	3.46

Five empirical models generated from continuum and line profile observations are compared to four theoretical models in Table IX(S4). Finally, including small corrections for the Hydrogen/metal ratio (Log A = 3. 95, B = 0. 18), an effect which raises the temperature by 30K° - 60K°, the photospheric model adopted by Minnaert[6] is presented in Table X(S4). Here the scale of heights is so adjusted that zero corresponds to the limb, which in the scale of optical depths is located near

$\tau_{5000} = 0.003.$

Table IX (SY)

MODELS OF SOLAR PHOTOSPHERE, T (LOG P) (°K) (REF. 15)

	FROM OBSERVED CONTINUUM			THEORETICAL				FROM OBSERVED PROFILES		
LOG P	PEY-TURAUX	BAR-BIER	DE JA-GER V (1952)	LABS	ALLER AND PIERCE	MUNCH	MI-CHARD	DE JA-GER III (1948) AND VOIGT (1950)	DE JA-GER VII (1951)	ADOPTED
(-∞).....	.....	.....	.....	4400	.....	.....	.....	.....	.....	4400
4.00.....	.....	.....	.....	4670	.....	.....	.....	.....	.....	4730
4.40..	5060	.....	4860	4800	.....	4790	.....	.....	4580	4900
4.60.....	5220	.....	5030	4970	4930	4920	4780	.....	4850	5070
4.80..	5570	5320	5350	5265	5300	5150	5050	5400	5220	5390
5.00.....	6150	5830	5830	5830	5750	5580	5500	5800	5900	5870
5.10.....	.....	6300	6170	.....	5880	.6080.	.5850.	6190	6220	6240
5.20.....	.....	.....	.....	.....	6770	.....	.....	7140	6800	6990
5.30.....	.....	.....	.....	.....	.....	.....	.....	7770	7650	7750
5.35	.....	.....	.....	.....	.....	.....	.....	8020	7900	8000

Table X (S4)

ADOPTED PHOTOSPHERIC MODEL (REF. 15)

h (Km)	T (°K)	$\mathcal{R}$	log P	log Pe	ρ (Gm/Cm^3)	k_{5000}	τ_{5000}
+∞.....	4300	1.172	.....	.....	.....	.....	.....
+361.....	4330	1.164	2.00	8.26-10	0.04×10^{-8}	0.001	0
+148.....	4460	1.113	3.00	9.20-10	0.40	0.007	0.0001
- 65.....	4730	1.065	4.00	0.16	3.8	0.053	0.011
-155.....	4900	1.028	4.40	0.54	9.31	0.108	0.056
-202.....	5070	0.993	4.60	0.78	14.3	0.159	0.129
-251.....	5390	0.935	4.80	1.04	21.3	0.222	0.294
-304.....	5870	0.858	5.00	1.41	30.9	0.346	0.672
-332.....	6240	0.808	5.10	1.70	36.6	0.587	1.11
-363.....	6990	0.722	5.20	2.26	41.3	1.38	2.19
-398.....	7750	0.650	5.30	2.82	46.8	3.85	6.09
-418.....	8000	0.630	5.35	3.02	46.9	5.70	10.3

VI. CONCLUSIONS AND RECOMMENDATIONS

Most solar atmospheric models agree rather well with one another in the intermediate ranges of optical depth. They deviate from one another in the shallow layers ($\tau < 0.1$) and in the very deep layers ($\tau > 1.0$) where the limb darkening studies can give no information.

A major problem of earth-based observations is energy extinction caused by the chemical constituents in the earth's atmosphere. Water vapor and carbon dioxide molecules severely hamper accumulation of the infrared solar spectra. Similarly, shorter wavelength absorption occurs due to ozone produced in the earth's upper atmospheric strata. Deficiencies in the solar spectrum yield inaccurate elemental abundance data which in turn generate unreliable hydrogen-to-metal ratios.

Observations are also limited by the seeing conditions imposed by the atmosphere. The practical resolution limit of ground-based solar observation is about 0.5 second of arc, even though the instrument's diffraction limit is less than this. The promise of achieving greater resolution in space is supported by the successful Stratoscope experiments (high altitude balloons) in which results approached the diffraction limit of the telescopes in use.

In the development of a solar atmospheric model, we conclude that the damping constants are not sufficiently well known for strong lines. The use of weak lines for which the damping constants are unimportant and for which a simple theory of line-formation may be employed should be explored. This would further require the theoretical development of f-values for the weak elements.

Furthermore, refinements in the theory of solar atmospheric models, e. g., correction for the failure of the plane-parallel stratified atmospheric layer hypothesis, have to be investigated to include the influence of the chromosphere on lines of moderate to high intensity. Models of the solar chromosphere and corona become increasingly difficult, because the diversity of information is less in both quantity and quality. More observations and analyses are necessary so that solar atmospheric modeling may be extended to these regions.

VII. REFERENCES

1. Aller, L. H., "The Atmosphere of the Sun and Stars", Astrophysics, Ronald Press Company, New York, 1963, p. 129.

2. Wrubel, M. H., Handbuch Der Physik, Vol. LI, Springer-Verlag, Berlin, 1958, p. 6.

3. Aller, L.H., Op. Cit., p. 225.

4. Chandrasekhar, S., Radiative Transfer, Dover Publication, Inc., New York, 1960, p. 290.

5. Aller, L. H., Op. Cit., p. 118.

6. Minnaert, M., "The Photosphere", The Sun, Ed. Kuiper, G. P., University of Chicago Press, Chicago, 1953, p. 97.

7. Moore, C. E., "The Identification of Solar Lines", The Sun, Ed. Kuiper, G. P., University of Chicago Press, Chicago, 1953, p. 202.

8. Chandrasekhar, S. and Breen, F., Ap. J., 104, 1946, p. 430.

9. Menzel, D.H. and Pekeris, C. L., M.N., 96, 1936, p. 77.

10. Abell, G., Exploration of the Universe, Holt, Rinehart and Winston, New York, 1964, p. 372.

11. Pawsey, J. L. and Bracewell, R. N., Radio Astronomy, Oxford at the Clarendon Press, 1955, p. 111.

12. Aller, L.H., Op. Cit., p. 303.

13. A. Unsold and V. Weidmann: Vistas in Astronomy, Vol. I, London: Pergamon Press, 1955, p. 249.

14. Aller, L. H., Op. Cit., p. 210.

15. Minnaert, M., Op. Cit., p. 129.

APPENDIX S-5

ENGINEERING CONSIDERATIONS FOR A SOLAR PROBE

By Robert O. Breault

The object of this appendix is to compile data on the feasibility of a solar probe mission. An eclectic survey was conducted with the aim of determining the state of the art in technological advances, in an attempt to avoid research redundancy and bring into focus the limitations and constraints that presently avail. Trajectory and orbital transfer, weight versus booster availability, power consumption, path and attitude control systems, perturbations, effects of solar pressure, solar thermal radiation and communications are discussed. Numerical and graphical information is presented throughout the appendix. Evaluation, but not prediction, is attempted by the author.

I. ELLIPTICAL ORBITS

Feasibility calculations for solar probes would have to account for the elliptical orbit of the earth, obliquity of the ecliptic, gravitational effects of the Sun, drag forces, solar radiation pressure and possible pertubation effects of Mars, Venus and Mercury. Studies have shown[1] that the optimum periods of time in which interplanetary orbital transfer requires a minimum energy expenditure, are early January, with the earth at perihelion, and early July, when the earth is at aphelion.

In order to overcome the potential energy well of the earth an escape velocity of 11.2 km/sec must be attained. This may be obtained by a multistage space vehicle that has the advantage over a single stage vehicle of a better mass ratio. A solar probe in a parking orbit about the earth must now be impulsed, so that it can be put on a ballistic trajectory. The transfer would be accomplished by a Hohmann transfer ellipse. The Hohmann ellipse

has the special property of requiring the least energy for a two-impulse transfer. According to Edelbaum[4], "such maneuvers also minimize fuel consumption for transfer from a circular orbit to coplanar orbits which satisfy any desired functional relationship between energy and angular momentum and which have any desired orientation. Examples of these functional relationships are fixed energy (or major axis or period), fixed angular momentum (or latus rectum), fixed apogee or perigee radii, and fixed eccentricity and major axis."

Although other trajectories requiring less energy may be found, more impulses are required.

Dugan[1] has shown that solar probe orbits become highly eccentric as the perihelion distance approaches the radius of the sun. In addition, the period of the solar-probe orbit is essentially linear with respect to the perihelion distance out to 0. 375 A. U. (Mercury perihelion is 0.315 A. U.). For a given perihelion distance, the period for favorable launch is about one week longer for summer than for winter launchings since the Earth aphelion occurs during the summer.

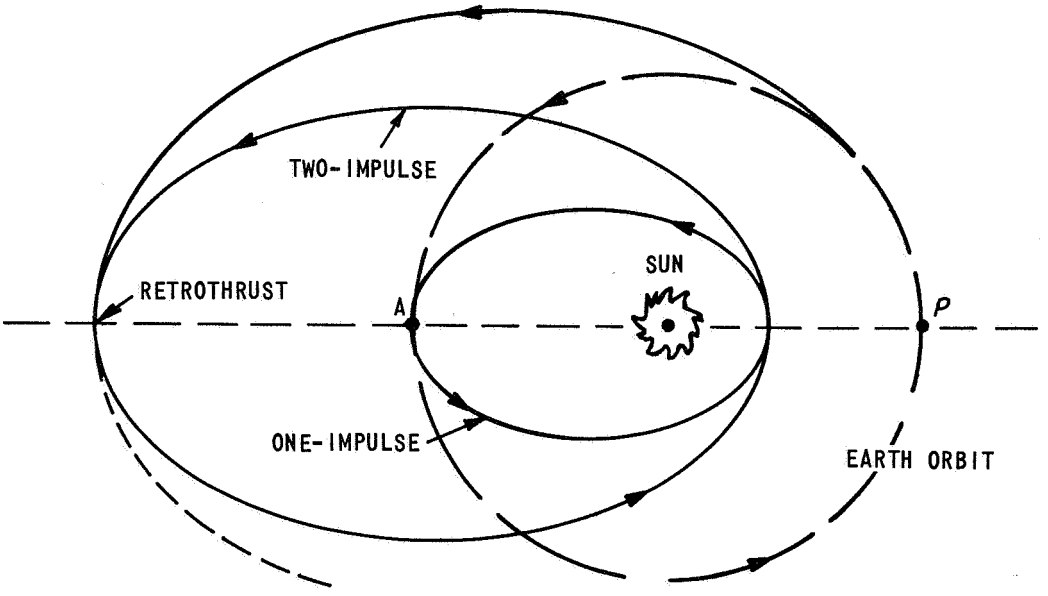


Figure 1(S5) ONE- AND TWO-IMPULSE METHODS. (REF. 1)

Elliptical orbital transfer may be affected by either the single-impulse or a two-impulse method. For Earth-launched probes with a residual velocity in the same direction as the orbital velocity of the earth, the two-impulse method would require a retrothrust at the probes aphelion [Figure 1 (S5)].

Although closer orbits to the sun may be obtained with a two-impulse technique, greater time and additional complexity are inherent factors. For a single-impulse, launches in the direction opposite to the orbital velocity of the earth will give a coplanar, elliptical solar-probe orbit. It should be pointed out that the two-impulse method offers no velocity advantages for perihelion distances greater than 0.22 A. U. [Figure 2 (S5)]. In Figure 2 (S5), r_a is the retrothrust distance from earth launch.

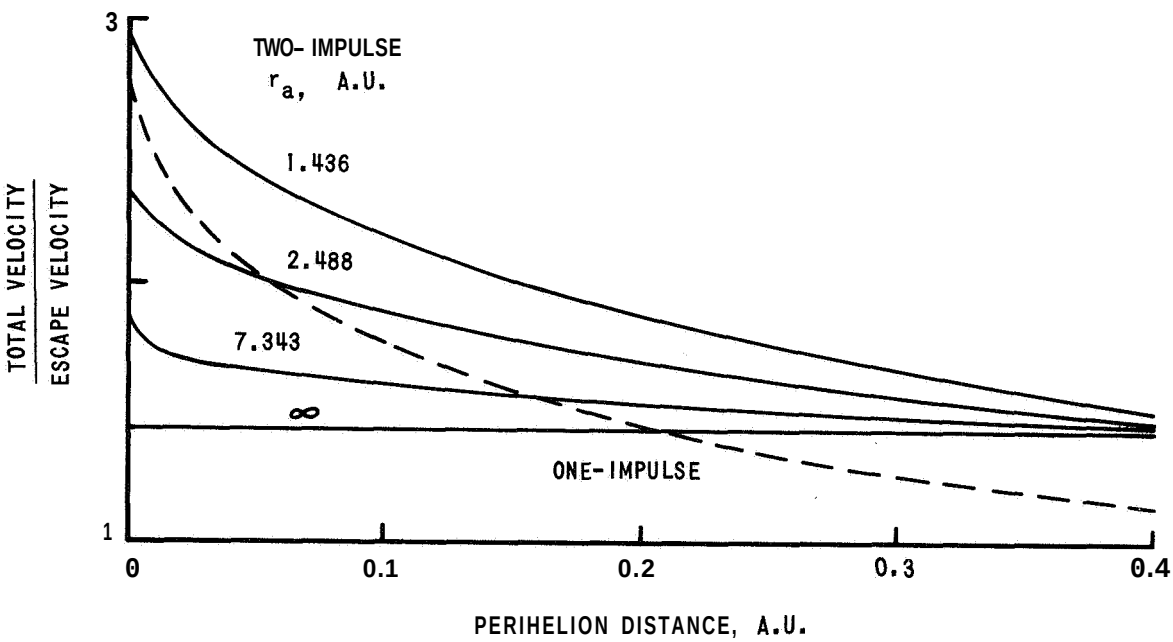


Figure 2(S5) VELOCITY REQUIREMENTS (REF. 2)

Results calculated by Hall, Nothwang and Hornby[2] have stated that although the two-impulse method has the advantage over the one-impulse method for distances less than 0. 2A. U., the time to reach perihelion is extremely long. For a perihelion distance of 0.07 A. U., a total velocity equal to twice

the escape velocity for the Earth would be required with either the one-impulse or two-impulse method and additionally a retrothrust at 2.488 A. U. for the two-impulse technique. For a single thrust it would take 2-1/2 months to reach perihelion, whereas it would take 4 years to reach perihelion by the two-impulse method.

For an orbit having a 0. 03 A. U. perihelion, a 2% velocity error and 1° angular error at burnout would result in only a 5% change in perihelion.

To insure continuity in communications, maximum optical resolution and photogrammetry, solar-probe launchings should occur during the years of the "quiet sun", i. e., 1974, 1985 and 1996.

II. BOOSTER CHARACTERISTICS

The question now arises as to whether or not there is present, or projected, technology capable of producing the velocities required to obtain the orbits discussed in the previous section.

Two operational boosters, the Atlas Centaur and Saturn C-1, will be discussed. A Post-Saturn (NOVA) with a predicted 12×10^6 lb. thrust and a heat transfer nuclear rocket in which the conventional combustion chamber of the liquid chemical rocket is replaced by a nuclear reaction, i. e., heat from the fissioning uranium in the fuel elements is transferred to the working fluid (hydrogen propellant) by conduction, convection and radiation, are scheduled to be operational by 1970 and 1975, respectively.

Values for the booster characteristics have been determined[2] and are shown below. The Atlas Centaur has an additional solid propellant upper stage (X-259) while the Saturn C-1 has a liquid H_2-F_2 upper stage.

TABLE I (S5)

BOOSTER DATA

Booster	Weight (lbs.)	Thrust (lbs.)	Altitude (ft.)	Velocity (ft/sec)
Atlas Centaur +X-259	293,000	367, 000	0	0
1st Stage	3,100	19, 000	600,000	34,200
Payload (Post Sep)	250	-	620,000	50,200
Saturn C-1 t H_2-F_2	985,000	1, 500,000	0	0
1st Stage	22,000	10,000	500,000	27,600
2nd Stage	8,000	4,000	2, 150, 000	38,600
3rd Stage	1,750	1,000	18, 000, 000	51, 000
Payload (Post Sep)	250	-	46, 250, 000*	66, 900

* 8, 760 miles

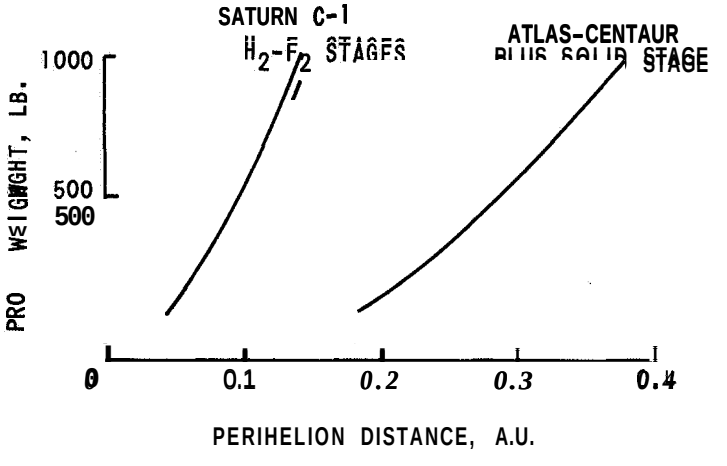


Figure 3(S5) PERFORMANCE OF SOLAR PROBE (REF. 2)

From Figure 3(S5), perihelion distances of 0. 066 A. U. and 0.24 A. U., for probe weights of 250 lbs., may be obtained by the Saturn C-1 and Atlas Centaur boosters, respectively. It should be noted that an Atlas Agena B having a solid rocket on top could achieve a perihelion distance of 0.012 A. U.

Examining the Atlas Centaur X-259 booster, the period of the orbit achieved for a probe weight of 250 lb. would be 6 months. A Sun-Earth axis 6 month orbit factually implies a 65 ± 5 day, 118 ± 5 day and 183 ± 5 day, after launch breakdown in communication, because the solar probe's ground antenna would be pointing directly toward the Sun [Figure 4 (S5)].

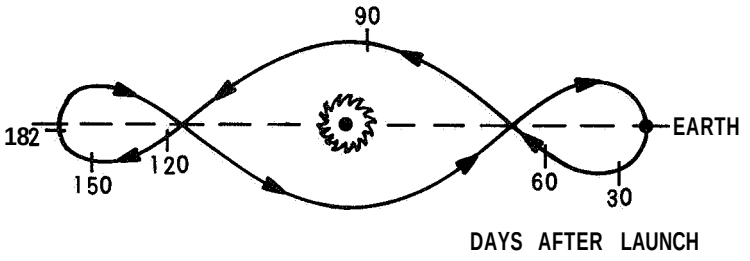


Figure 4(S5) SIX-MONTH ORBIT RELATIVE TO SUN-EARTH-AXIS (REF. 2)

After the first complete orbit the solar probe will be located in the vicinity of the earth's orbit, but the probe's observations will be unaffected by geo-magnetic fields and the blanketing effect of the atmosphere. It is also noted that the probe's heliocentric coordinates from 3 months to 8 months following launch, when the probe would be in the line of sight to the Earth, would be in the Sun's corona where low frequency radio transmission experiments would be recommended (see Communications).

Figure 5 (S5) shows the residual velocity (ft/sec) and orbital inclination (deg) required for various perihelion distances. It can be seen that to achieve a probe orbital perihelion of 0. 1 **A. U.** (9, 000, 000 miles = 21 Solar Radii), a residual velocity of 62, 000 ft/sec with an orbital inclination of 25° is indicated. Since the Sun's equator is inclined 7° to the ecliptic, experimental measurements as much as ±32° from the equatorial plane may be achieved.

Ideal burnout velocities corresponding to desired perihelion distances (Solar Radii) are shown in Figure 6(S5).

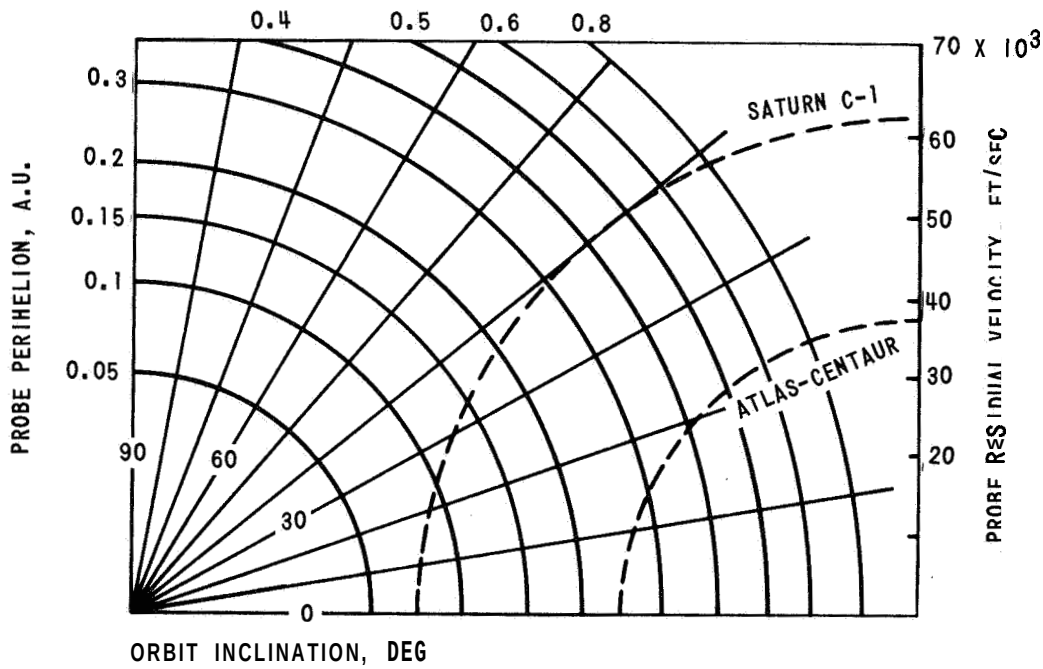


Figure 5(S5) VELOCITY REQUIREMENTS FOR ORBITS ABOUT SUN AND OUT OF ECLIPTIC PLANE. (REF. 2)

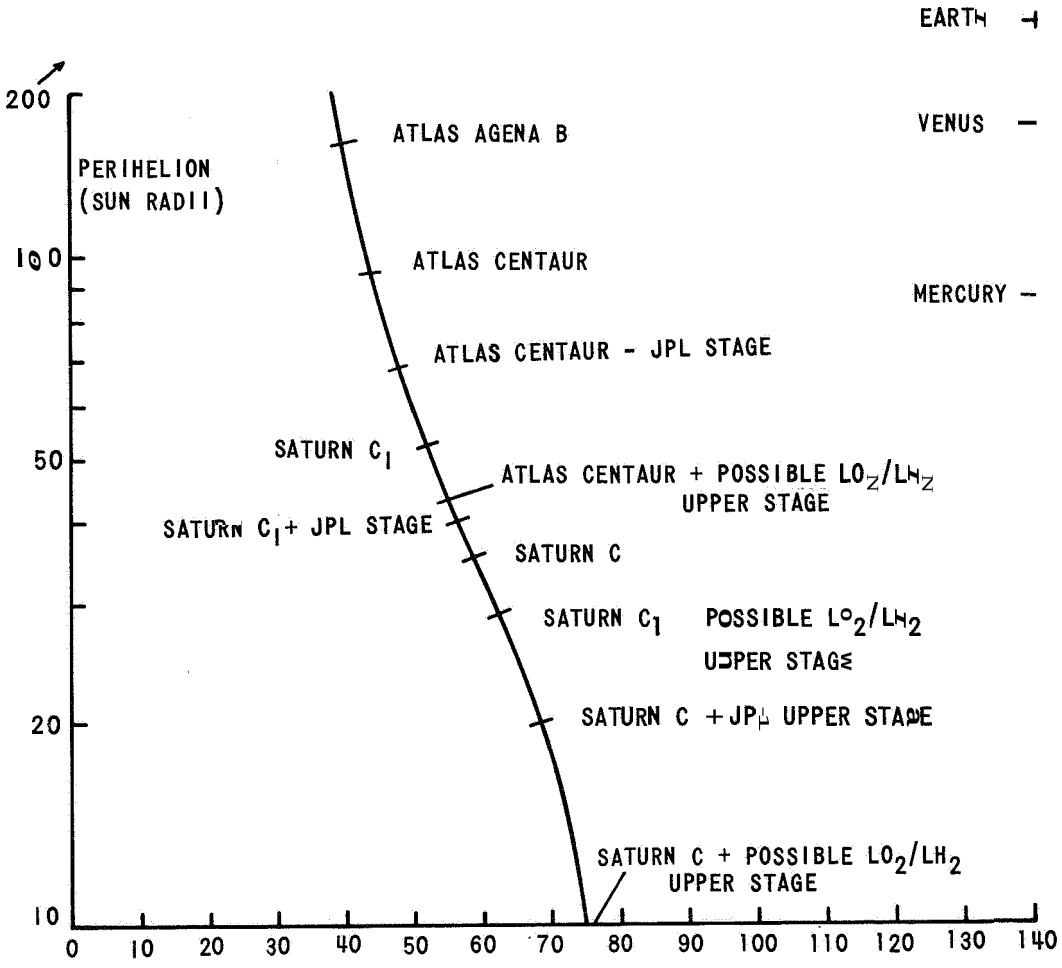


Figure 6 (S5) IDEAL BURNOUT VELOCITY - KILOFEET PER SECOND MINUS 400 FT. PER SEC. (REF. 3)

III. POWER SUPPLY

The necessity of power supplies is to transmit data and operate equipment. The amount of power required will be dictated by a variety of factors[10] which include, the frequency bandwidth, the limit of channel capacity, carrier frequency, signal-to-noise ratio, distance from transmitter to receiver and experimentation. To determine an order of magnitude of the power required, consider the modified Mariner II Venus fly-by. In this example, a carrier frequency of 960 Mc with a transmitter power of 3 watts successfully sent 8.33 bits per second up to a distance of 54, 000, 000 miles (0.58 A. U.) from earth. An S-band range transponder transmitter, having a carrier frequency of 2, 295 Mc and requiring 10 watts of power was also used.

The amount of power necessary for experimentation in a solar probe would be on the order of 20 watts*. Other power requirements for experimentation would be comparatively negligible since attitude control would be provided by the sun's gravitation and radiation fields, and temperature control of the instrumentation actuated by a liquid-gas system.

Solar power supplies may be of two classes: those designed to be maintained with a fixed geometric relationship to the sun--termed oriented; and those which are capable of satisfactory operation in any position relative to the sun--termed unoriented. The latter require more solar cells but do not require attitude control.

---

* Orbiting Geophysical Satellite, weights 150 lbs. , draws 50 watts (total), Ranger deep space probe, weights 57 lbs., draws 8 watts (total).

A feasible source for secondary power is solar cells. Since the main adverse characteristic of solar cells is their temperature sensitivity, arrangements where the solar cells are inclined to the sun's rays have been proposed[2]. At a distance of less than 1/2 A. U., the panels are at an angle of 10° (angle between panel and probe axis). The angle becomes 70° for distances between 1/2 A.U. and 1 A.U. At the small angle (10°) the solar probe receives less solar radiation per unit area and its temperature is also less. The power curve in Figure 7 (S5) represents the requirements for the instrumentation and communication when the probe is at the opposite side of the Sun from the Earth. The required communication

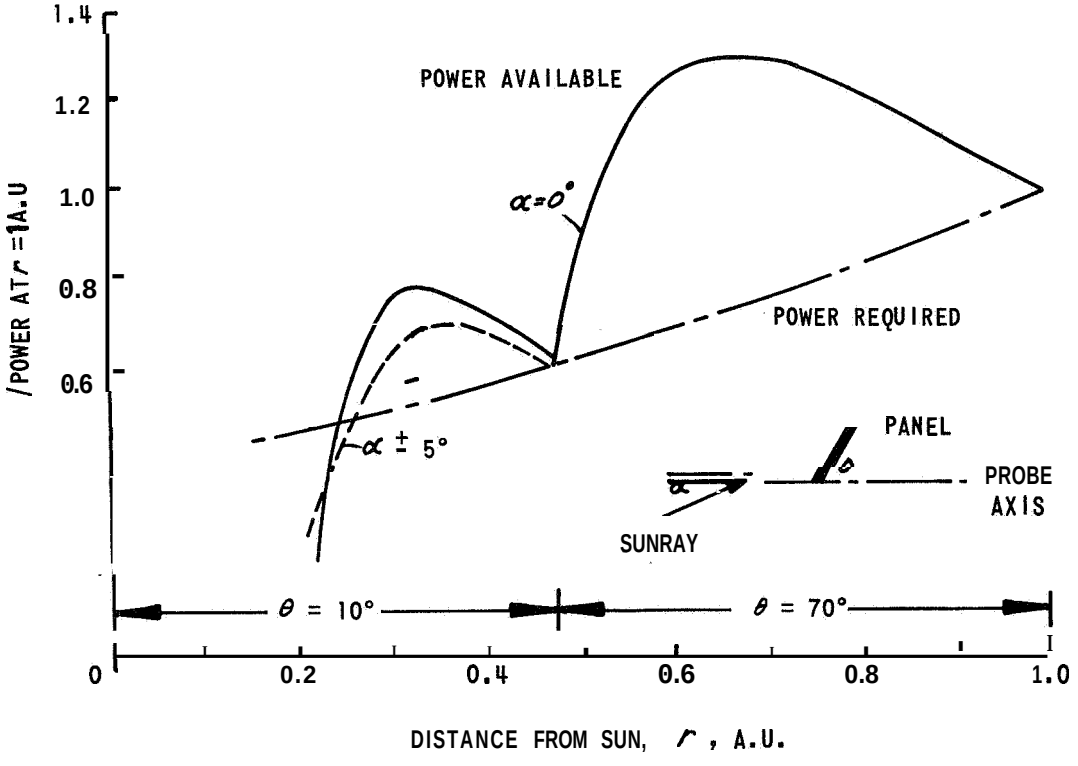


Figure 7 (S5) POWER OF DEFLECTABLE SOLAR CELL PANEL. (REF. 2)

power is at a maximum when the probe is 2 A.U. from the earth. Performance values indicated that specific power of 2-1/2 watts per lb. and between 2 and 3 watts per sq. ft. of surface area have been used for conditions at 1 A. U.

An upper limit specific power of 10 watts per lb. of weight €or efficient on-board power systems is predicted.

Solar Cells

Orbital conditions of insolance (received solar radiation) and eclipse, the period of illumination (100% to 50%), the degree and variation of orientation and incidence angle of irradiation, as well as solar intensity, indicate that the properties of silicon solar cells constrain the range, for useful insolance, from 0.43 A. U. to 2 A. U. from the sun[5]. Shadowing effects due to improper vehicle design could reduce the value of the output voltage, and should be taken into consideration. Calculations by other workers in the field show useful ranges closer to the sun [see Figure 8(S5)].

Usually the primary load is an electrochemical storage battery whose efficient operating life-time is affected by the depth of the charge-discharge cycle, overcharging, solar cell voltage and, of course, temperature. For an ideal cell, the e-i characteristic is almost rectangular, The nearly constant current characteristic on the short circuit side of maximum power is well suited to the ideal charge condition for a battery, whereas the nearly constant voltage characteristic on the open circuit voltage side of maximum power is well suited to terminal voltage limitation as the fully charged condition of the battery is approached. This technique may be used only when the solar panel temperature is nearly constant and accurately known[6].

At usual operating temperatures individual silicon solar cells provide 0.3 to 0.5 volt at a power level of a few mw/cm^2. The e-i characteristics of a typical p-n silicon solar cell at various illumination levels have been examined and show that except at very low levels, current is approximately proportional to insolance and voltage is essentially independent of insolance. In addition, the dynamic internal impedance of the cell is approximately inversely proportional to the illumination level. The effect of temperature on the e-i characteristics of a typical p-n silicon cell shows that the effect on current is relatively small, but power, voltage, and internal impedance decrease rapidly with increasing temperature,

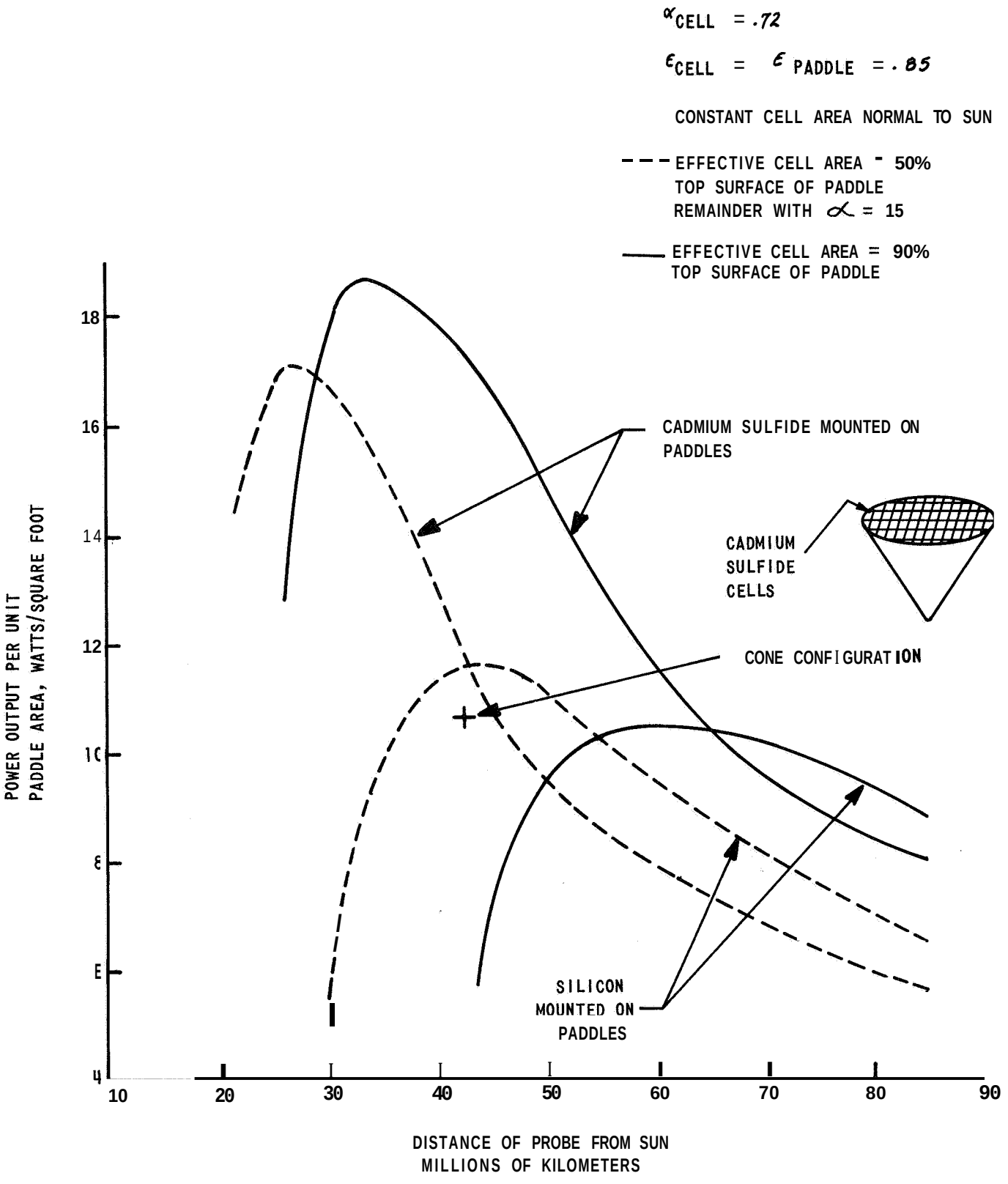


Figure 8 (S5) SOLAR CELL OUTPUT (REF. 3)

The heat input to a solar array in space is the summation of absorbed direct solar energy, albedo and thermal energy from nearby auxiliary (vehicle) surfaces, internal power losses, and conduction from other portions of the vehicle. The energy output is the summation of thermal radiation to space and nearby surfaces, conduction to other portions of the vehicle, and converted electrical energy. The equilibrium temperature is that temperature at which the input and output energies are equal. The equilibrium temperature for a space vehicle in a near earth orbit, using uncoated cells, for the solar irradiance of 140 mw/cm^2 is 86°C, with a cell efficiency of 6.7% and a panel efficiency of 5.7%. Recently, multi-layer interference films of dielectrics such as TiO_2, ZnS, MgF_2 and SiO_2 have increased emissivity from 0.31 to 0.87 on a 50°C black body.

The chief environmental dangers to solar cells in space lie in micro-meteoroid bombardment and Van Allen Radiation. Van Allen Radiation damage results primarily from high energy proton and electron bombardment. A summary is given in Table II(S5).

Protection from radiation damage may be provided by glass covers from 10 to 70 mil range, i.e., 3 mev protons are stopped with 3.5 mils of glass, 12 mev protons require 38 mils of glass and 800 kev electrons require 65 mils.

TABLE II (S5)

EXPERIMENTAL DATA
REDUCTION OF SOLAR CELL EFFICIENCY
DUE TO PARTICLE BOMBARDMENT

Source	Particle	Energy (mev)	Flux (part/cm^2)	Percent Reduction
Lockheed	Proton	3.0	5×10^9	25
		12.0	2×10^{11}	25
		12.0	10^{12}	50
STL	Electron	0.5	5×10^{13}	25
RCA	Proton	17.6	10^{15}	50
	Electron	1.7	5×10^{14}	50

The use of optically treated glass slides, mounted on the face of each individual cell, is primarily for defense against excessive thermal solar radiation. The glass slides are cemented to the solar cells thus lowering the equilibrium temperature as determined by thermal radiation balance so that a supplemental radiating surface is not required. Heat is transferred by a combination of conduction and radiation paths throughout the solar power supply ring so that those panels which are illuminated are partially cooled by radiation from the panels which are not illuminated. The glass slides are equipped with an ultraviolet absorption filter and an anti-reflection coating on the inner face. Since the cells are not responsive to the ultraviolet, little loss is introduced by the filter which serves the purpose of protecting the epoxy cement used to bond the glass to the cells. The glass slides also afford some mechanical protection to the thin active layer on the solar cell face, although their main function is temperature control.

A second, but less known, method for obtaining secondary power system output levels up to 20 kilowatts is the use of solar concentrating mirrors. Collectors of low specific weight, less than 0.3 lb/ft^2, having efficiencies above 70% at absorber temperatures of 3000°F have been evaluated[7] and seem eminently feasible.

IV. PATH AND ATTITUDE CONTROL SYSTEMS

The purpose of the path control system is to orient the thrust vector in a specified direction. This is accomplished by control of the orientation or rotational motion about the center of mass during periods of unpowered flights. If the disturbing torques are small, (separation from the launch booster, magnetic fields, plasmas, energetic particles and micro-meteorites of space, non-uniform equipment rotation and antenna radiation pressure), the control torque is very small and solar radiation pressure may be utilized functionally.

A common method used for powered flight (attitude control) is to hinge the main rocket thrust chambers[8]. Yaw, roll, and pitch control channels may be readily obtained. However, if the vehicle has only one main rocket nozzle, auxiliary jets are required for roll control.

The attitude-control functions may be divided into three categories; 1) those which require stabilization to an inertial or nearly inertial orientation reference, 2) those which require stabilization with respect to an Earth-centered or other body centered reference (one axis is toward the center of the body and another axis is in the forward direction in the plane of the orbit), 3) those which involve large angle slewing from one orientation to another. Table III (55) gives a partial listing of satellites and accompanying references.

The range of variation of quantitative attitude requirements may be divided into three categories; a precise range whose upper bound is probably a few arc minutes (for the OAO the requirement is 0.1 sec of arc with respect to an inertial reference), a "normal" range of 0.5° to several degrees, (a solar probe must be held to $\pm 2.5°$ of the desired direction about its longitudinal or roll axis because of antenna directionality [see Ref. 2]) and a low accuracy range beginning about 10°.

TABLE III(S5)
ATTITUDE CONTROL SYSTEMS

Satellite	Reference
Discoverer Series	Earth-centered, 3-axis control
Explorer Series	Spin stabilized, 1 axis inertially fixed
Mercury Series	Earth-centered, 3 axis control; also gross attitude changes
Nimbus	Earth-centered, 3 axis control
Orbiting Astronomical Observatory (OAO)	Inertial, 3 axis control; also gross attitude changes
Orbiting Geophysical Observatory (OGO)	Earth-centered, 3 axis control solar paddle control and gross attitude changes
Orbiting Solar Observatory (OSO)	Spin stabilized, 1 axis inertial with 1 axis toward the sun

V. PERTURBATIONS

It was mentioned, in Section I, that the period of the solar probe orbit is essentially linear with respect to the perihelion distance for a given launch point. In addition, the eccentricities of the solar probe orbits are assumed constant and, the orbits are essentially co-planar with the plane of the ecliptic. Projecting these conditions to their logical conclusion results in the possibility of the solar probe passing into the Earth's gravitational sphere of influence when an inferior conjunction occurs simultaneously with an aphelion passage,

The above premises become accurate and acceptable only when the assumption that the trajectories and orbits of the solar probes are not influenced by perturbations from the Earth's gravitational field, the Moon, Venus or Mercury, The earth's influence becomes negligible at about 0.016 **A.** U. from the earth. The largest perturbating effect would occur when the orbit of the solar probe crosses the orbital plane of the influential mass. Dugan[1] has heuristically argued that the assumptions remain valid and perturbating effects are negligible provided the time of launching is so selected that the distances of the probe relative to the Moon, Venus or Mercury, as the probe crosses these orbital planes, is sufficiently large.

VI. SOLAR RADIATION PRESSURE

Studies have shown[3] that the torque exerted on a space probe's collector due to solar radiation pressure is proportional to (a) the area of the collector, (b) distance of the collector from the probe's center of gravity, (c) the inverse square of the probe's distance from the sun, (d) and the sine of the angular deviation of the collector from the sun. This can be used for attitude control.

The use of solar radiation pressure for propulsion of space vehicles has been investigated (for fly-by probe)[1, 9] and in view of the uncertainties associated with solar sailing, solar radiation pressure is not considered a means of propulsion[9].

VII. SOLAR THERMAL RADIATION

The minimum perihelion distance from the Sun, for a solar probe, would be dictated by the highest temperature at which the probe and its instrumentation would operate efficiently. Although the coronal temperature is calculated to be 1,000,000°K, the extremely long mean free path of the solar atmospheric constituents indicate that heating from coronal sources may be negligible. Direct thermal radiation and dissipated internal telemetry heating appear to be the major considerations controlling upper boundary temperature constraints.

Dugan[1], treating the Sun as a black body, has calculated absorptivity and emissivity curves for two spherical solar probes: one of silver, the other of aluminum. He assumed that the following rule for metal holds - the absorptivity of a metal surface at temperature T_1 with respect to radiation from a black body at T_2 is equal to the emissivity at a temperature $(T_1 T_2)^{1/2}$. It was found that the ratio, absorptivity/emissivity, decreases with increasing temperature of the metal surface, thereby tending to reduce the variation of temperature with distance from the sun. For a spherical probe, of uniform surface temperature, a 0.6 A. U. perihelion would expose the payload to a temperature of 600°K. A perihelion distance of this order of magnitude would have little value. For a spherical probe, with one side continuously oriented toward the Sun, approaches closer to the Sun are feasible. Optimum perihelion distance may be achieved with a solar probe geometrically constructed as a hemisphere followed by a right conical frustum. A distance of 0.05 A. U. appears obtainable where the rate of heat absorbed by the probe from the sun is 1400 BTU/min (25 kilowatts) for a probe of radius 1 foot. The probe's equilibrium temperature would be 1000°K (Ref. 1).

The temperature achieved for a multi-shielded probe are given in Figure 9 (S5).

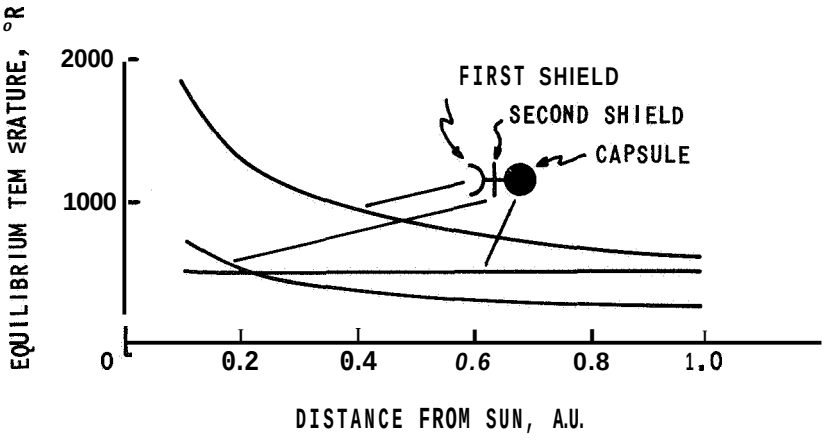


Figure 9 (S5) TEMPERATURES FOR CAPSULE AND INDEPENDENT HEAT SHIELD. (REF. 2)

Radiation is the only means for removing heat in outer space. It appears feasible to utilize a thermionic converter whose anode (cold) temperature may be as high as 900°K. Waste heat from the converter may be conducted from the anode through a metallic path to the radiating surface. For a solar probe whose peak power is about 1/2 kw radiating fins may be employed.

VIII. COMMUNICATIONS

The field of spaceprobe communication is broad and extensive. To avoid possible redundancy, only a general method for transmission and reception will be discussed.

It has been suggested[**]), that a solar probe experiment could be devised utilizing a low frequency receiver and a second radio transmitter having a frequency considerably below the communication frequency. A radio frequency of approximately 100 megacycles would be transmitted from the probe to the Earth in addition to the higher communication frequency.

The electron density in the corona of the Sun, could then be determined from the amplitude and phase change between these two waves. If the low frequency wave is plane polarized and the plane of polarization is slowly rotated, the component of the magnetic field perpendicular to the direction of propagation of the wave may be inferred.

During the time that both frequencies are transmitted, information would not be transmitted from the communication system and its bandwidth would be reduced to less than 10 cycles per second, thereby reducing the amount of power required at the higher frequency, The surplus power would then be used to drive the lower frequency transmitter. It is estimated that the power required for the low frequency transmitter with a bandwidth of about 1 cycle per second plus the power required for the communication frequency with its reduced bandwidth would be approximately equal to that required for the communication system to transmit information at the bandwidth required by the information rate. Table IV (S5) summarizes suggested data.

TABLE IV (S5)

SOLAR PROBE **DATA**

Frequency	2.3 kmc
Transmission Distance	2 **A.** U.
Ground Antenna	85 ft. diameter
Probe Antenna	**3** ft. diameter
Receiver Temperature	100°K
Bandwidths	60, 100, 225 cps
Signal-to-Noise Ratio	9.5, 10.5, 11.5 db
Power Upper Value	100, 210, 610 watts
Power Lower Value	60, 130, 370 watts
Noise Losses	9 db

IX. CONCLUSIONS

From the considerations presented in this memorandum, we see that current literature[''] indicates a perihelion distance of 0.05 A.U. for solar probes may be attainable. It should be noted that the inference was reached irrespective of the experimental instrumentation carried on the probe. The adverse effects of the solar environment on instrumentation have not been studied and investigations along these lines would have to be conducted before a practically achievable perihelion distance can be defined.

X. REFERENCES

1. Dugan, D. W., "A Preliminary Study of a Solar Probe Mission", NASA Technical Note D-783, Ames Research Center, April 1961.

2. Hall, C.F., Nothwang, G. J. and Hornby, H., "A Feasibility Study of Solar Probes", Institute of the Aerospace Sciences Paper No, 62-21.

3. Dickstein, D. H., "A Solar Probe", IAS Paper 61-179-1873.

4. Edelbaum, T. N., "Some Extensions of the Hohmann Transfer Maneuver", ARS Journal, Vol. 29, No. 11, November 1959.

5. Evans, A. E., et al, "Solar Panel Design Considerations", ARS Space Power Systems, Vol. 4, 1960, pp. 79-129.

6. Wolf, M., "Limitations and Possibilities for Improvement of Photovoltaic Solar Energy Converters", Proc of IRE, Vol. 7, July 1960, pp. 1246-1263.

7. McClelland, D. H., "Solar Concentrators for High Temperature Space Power Systems", ARS Space Power Systems, Vol. No. 4, 1960, pp. 129-152.

8. Seifert, H., "Space Technology", New York, Wiley & Sons, pp. 24-22, 24-23.

9. Garwin, R. L., "Solar Sailing-A Practical Method of Propulsion Within the Solar System", ARS Jet Propulsion, Vol. 28, No. 3, March 1958, pp. 188.

10. Campbell, Charles E., "Project TECH TOP, Study of Lunar, Planetary and Solar Topography", Phase I Report, Cornell Aeronautical Laboratory, Inc., Report No. VC-2104-D-1, August 1966.

APPENDIX S-6

SOME CONSIDERATIONS OF THE USE OF COOPERATING VEHICLES TO OBTAIN TOPOGRAPHIC INFORMATION IN SPACE

By John D. Gallatin

I. INTRODUCTION

This appendix discusses some potential applications of co-operating vehicles for space missions designed to obtain topographic information. Consideration is given to situations in which one of the vehicles acts as a relay station for the other, to situations in which co-operating vehicles are to be used because time is of the essence, and to cases in which neither of these apply. The advantages and disadvantages of the potential applications are discussed and conclusions are presented.

II. RELAY STATIONS

Acquisition of topographic information by means of cameras or reconnaissance radars requires that the camera or radar antenna be more or less pointed toward the surface under investigation. However, transmission of information is facilitated by the use of a high gain transmitting antenna, if the antenna is kept pointed toward the receiving station (on earth). This simultaneous pointing requirement can be satisfied if the plane of the orbit of the satellite is normal to the line of sight from the satellite to the earth. Such orientation may not be desirable because (a) additional energy may be required to set up the orbit in this manner, and (b) corrections would be required to overcome the algebraic difference between the regression of nodes and the relative revolution of the earth and satellite's primary. Furthermore, it may not be possible because the plane normal to the earth-to-satellite line may not be a plane from which acceptable imagery can be obtained.

Simultaneous pointing can also be satisfied by moving the antenna with respect to the camera or radar. This requires the use of pivot bearings which can withstand the space environment and would likely interfere with sensor operation, particularly in the case of radar. A really high gain radar antenna might be several times as large unfolded as the rest of the satellite. The interference is caused by (a) shadowing and (b) induced angular motion.

These difficulties could be avoided by use of two separate satellites, one to carry the camera or radar, and the other to carry the transmitting antenna. Both would be placed in the same orbit, one somewhat ahead of the other. Data transmission between the two satellites would be accomplished by omnidirectional antennas which could be used because of the short distance involved.

If it becomes possible, and is considered worthwhile, to monitor the sun by means of a space probe in a nearly circular orbit close to the sun, a relay satellite may be required. The desirability of such a satellite would depend on the gain pattern of the earth-located final receiving antenna, the power transmission and gain pattern of the antenna at the primary satellite, and the feasibility of considerable data storage at the primary satellite. In the remainder of this section, the method of computing the approximate data storage interval required is shown; if this storage interval is believed to be too long, a relay satellite is required. The parameters are chosen only to illustrate the computation and are not necessarily representative of any particular time period.

The minimum elongation at which communication can occur depends, as noted before, on the transmitting characteristics of the vehicle and the characteristics of the receiving antenna. If it is assumed that transmission is possible when the vehicle is more than 1" from the edge of the visible sun (this figure is to be determined from inspection of antenna gain pattern, etc.) then it may be seen by trigonometry that the minimum elongation required is 3.3×10^6 km. This is the sum of 0.70×10^6 km solar radius and 2.60×10^6 km corresponding to 1". Then, for transmission to occur, the

angle between the vehicle-sun and earth-sun Lines must be greater than arc sin (3.3×10^6 km/r) where r = radius of orbit. If r = 7.45×10^6km, or 0.05 A. U., then it is seen that the angle is 26.3" and transmission will not be possible during 29.2% of the orbit. Continuing with the 0.05 A. U. orbit, and using Kepler's third law, the sidereal period may be found to be 4.185 days. This results, assuming a prograde orbit, in a synodic period of 4.75 days. Because there are two intervals during which data must be stored, the storage interval is equal to 1/2 x 29.2% x 4.75 days, or about 16-2/3 hours. If it is found that data accumulated over 16-2/3 hours could not be stored at the temperatures involved, a relay vehicle leading or lagging the primary vehicle would permit continuous transmission. Also, such a relay could facilitate reacquisition after eclipses and transits.

III. COOPERATING VEHICLES, BOTH WITH SENSORS, WHEN TIME IS IMPORTANT

The motivation for the use of cooperating vehicles, both with sensors, is the desire to obtain three-dimensional topographic information either on transitory phenomena, or in real time. Both of these must be qualified, however. Use of cooperating vehicles to obtain information on transitory phenomena is required only if it is not possible or convenient to have a single vehicle with a three-dimensional capability, e.g., combination of radar and camera, or to have a single vehicle move through a sufficient distance during a time short compared to the time scale of the phenomenon of interest. Similarly, two vehicles are required for real time data presentation only if the single vehicle with three-dimensional capability is undesirable and the demand is for information in less time than would be required for a single vehicle to obtain three-dimensional information by use of two stations. No requirements for cooperating vehicles to extract real time information are obvious to this writer; however, there are interesting transitory phenomena associated with the Sun and Jupiter. Application of cooperating vehicles to the investigation of these phenomena is discussed below.

Because of the short time scale of some solar phenomena, especially granularities, and problems in obtaining orbits close to the sun, use of cooperating vehicles to obtain three-dimensional information of solar phenomena appears advantageous. The problems of not using cooperating vehicles are the large amount of energy required to obtain close orbits and the thermal environment of close orbits. The three-dimensional information could be obtained by combining images obtained by a vehicle with those obtained on earth. However if high resolution is required, such as if granularities are to be observed, better imagery may be obtained by using two vehicles so all images are obtained beyond the degrading effects of the earth's atmosphere. Appendix S-6A contains a derivation showing the lack of feasibility of obtaining three-dimensional information of granularities with a single probe.

The use of cooperating vehicles to obtain three-dimensional information of transient phenomena on Jupiter does not appear to be required. Jupiter has a short rotation period of about 10 hours, and some of the interesting phenomena seem to change slowly, so that it may be possible to obtain stereo pairs from earth-based telescopes. If such pairs do not yield the required depth resolution, then it may be necessary to employ a single vehicle. It should be noted, in contrast to the sun, that for Jupiter: (1) a retrograde orbit is about as easily achieved as a prograde orbit, and hence the rotation of Jupiter may aid in developing the total angular velocity, and (2) there are no thermal limitations on low orbits.

IV. COOPERATING VEHICLES FOR PLANETARY MAPPING

The use of cooperating vehicles, both with sensors, for mapping of planets and satellites without atmospheres does not appear to offer any advantages.

As discussed in the Interim Report (Reference 1), vertical errors in stereophotogrammetry are minimized when the overlap is near 50%. Thus, local mapping should be done from pairs taken from the same orbit, i. e., using forward lap rather than side lap, as forward lap may be controlled to be near 50%. Two vehicles, each supplying separate plotters, are not included in the category of cooperating vehicles.

A discussion of the use of cooperating vehicles for radar mapping of planetary topography must be very general because: 1) the motivation for radar is to penetrate atmospheres, 2) not much is known about these atmospheres, and 3) the characteristics of these atmospheres influence mission parameters greatly. The following paragraphs list some ideas to be considered as atmospheric data become available.

On the basis of atmospheric data, information rate available to return data, required detail in the radar image, the rotation of the planet, the length of time alloted to the mission and the area to be covered, it should be possible to determine the operating altitude, width of covered track for each vehicle, and number of vehicles required. The comparison of altitude to track width, together with estimated stability, will indicate whether it would be more advantageous to map the surface by obtaining *two* side looking radar views from opposite directions (as seen at the surface) or to use one side looking and one squinted side looking view. By squinted is meant a side looking radar with an antenna beam not perpendicular to the intended direction of motion. A combination of two radars, one squinted forward and one backward, is also possible. Combinations including squinted radars are to be preferred if atmospheric conditions require a large altitude compared to track width, and permit very good stability. If this latter combination is chosen, then each vehicle is an independent source of three-dimensional information and multiple vehicles may be said to be cooperating only in a very general sense.

If a polar orbit is not planned and if the use of one of the squint combinations discussed above is not favored by conditions, then the three-dimensional information may be obtained by use of a vehicle with radars looking to both sides, or by use of two vehicles each with a radar looking to one side. Study of error geometry reveals that if two similar vehicles are used on essentially parallel paths, then errors are minimized if one of the vehicles follows (as seen from the surface) the path of the other. This is most easily visualized if one lags the other by a whole number of orbits; however, the same imagery would be obtained regardless of lag. Within the constraint of similar vehicles and parallel paths, it may then be seen

that the only advantage accruing from the use of cooperating vehicles for acquiring topographic data, is that the distance between them may be measured to obtain additional topographic control. This measurement would be of greatest value if the radar on both vehicles were viewing the same part of the surface when the measurement was made. To synchronize and orient two satellites to do this is expected to be very difficult.

If polar orbits are contemplated, then a third possibility exists. The reverse view may be obtained by operation on both north and south bound portions of the orbit, using one vehicle with a radar looking to only one side.

V. CONCLUSION

The above study has revealed only the following applications of co-operating vehicles to the collection of topographic data: (1) as relay stations and, (2) in the collection of solar data.

VI. REFERENCE

1. Campbell, Charles E., "Project TECH TOP, Study of Lunar Planetary and Solar Topography", Phase I Report, CAL Report No. VC-2104-D-1, August 1966, p. 79.

APPENDIX S-6A

STEREO PAIRS OF THE SUN FROM CLOSE ORBITS

The difficulty of obtaining stereo pairs of fine detail on the sun, e. g., granularities, is that even with close orbits, the relative angular velocity of the probe would not be sufficiently large to permit observations from a large angle before significant changes occur. The small viewing angle then requires impossible angular resolution to obtain three-dimensional information. An example of this limitation is presented in this appendix.

The orbital period of a solar probe may be computed by Kepler's third law. Taking the year as the unit of time and the astronomical unit as the unit of distance, the equation needs no proportionality factors, i. e.

$$T^2 = a^3 \qquad \text{(S6A-1)}$$

where

T = period

a = semi-major axis.

For a probe in a circular orbit, the angular velocity with respect to sidereal space is

$$\dot{f} = 2\pi (r)^{-3/2} \qquad \text{(S6A-2a)}$$

For the general elliptical orbit, based on analysis not presented here,

$$\hat{f} = 2\pi r^{-3/2} \sqrt{\frac{2R}{r+R}} \qquad \text{(S6A-2b)}$$

where

f = angular velocity

$\hat{f}$ = angular velocity at perhelion

r = perihelion distance

A = aphelion distance.

To find the angular velocity with respect to the sun, the solar rotation must be accounted for. The net or total angular velocity is

$$\omega = \hat{f} \mp \frac{2\pi}{T_s} \qquad \text{(S6A-3)}$$

where

T_s = rotation period of sun.

the upper sign is to be used for prograde and the lower for retrograde orbits.

The angle $\omega \Delta T$, where ΔT is the time interval between observations, is the convergence angle. It is related to the resolvable depth by

$$\frac{\Delta D}{D} = \frac{\alpha}{\omega \Delta T} \qquad \text{(S6A-4)}$$

where

ΔD = resolved depth

D = distance to surface being investigated

α = resolved angle of system.

The above equations may be combined to yield

$$\alpha = \frac{\Delta D \Delta T 2\pi}{D}\left[r^{-3/2}\left(\frac{2R}{R+r}\right)^{1/2} \mp T_s^{-1}\right] \qquad \text{(S6A-5)}$$

For the following parameter values

$R = r$ = 0.05 A.U.

T_s = 27 days = 0.074 yr.

ΔT = 1 min = 1.90×10^{-6} yr

ΔD = 30 Km

D = 6.76×10^{9} meters (consistent with r = 0.05 **A. U.** and solar diameter of 1.39×10^{9} meters)

and a prograde orbit

α = 4.02×10^{-9} radian.

(retrograde: 5.5×10^{-9} radian).

The above values are, if anything, optimistic, because the photosphere is believed to be about 300 km deep, and significant relief could exist with a depth of less than 30 km, also, visible changes in granules occur in 1 minute. However, the required angular resolution of 4.02×10^{-9} radian is unreasonable. If this is taken as the Rayleigh resolution, then the required aperture for 500 $m\mu$ light is 98 meters. It may also be computed from the above formulas that even for a retrograde, parabolic ($R = \infty$) orbit with the same perhelion distance, the required aperture is 62-1/2 meters. Hence it is concluded that the use of a single vehicle is not a good method to obtain stereo pairs of granulation.

APPENDIX S-7

SOLAR INSTRUMENTATION

By Robert O. Breault

INTRODUCTION

The instrumentation utilized in the collection of astrophysical data has, in recent years, become more numerous and sophisticated. The increased variety of stellar work has generated an impetus for new technology. This appendix examines the functions and limits of instruments associated with Solar physics. These instruments include telescopes, both earth-based and balloon-borne, spectroscopes, gratings, spectrographs, spectroheliographs, photoheliographs, magnetographs, coronagraphs, radiation detectors and special filters. Additionally, construction criteria for optical surfaces are reviewed and geo-atmospheric limitations surveyed. The object is to present a broad, but comprehensive, foundation concerning data gathering instrumentation associated with the Sun.

I. TELESCOPES

Telescopes may be classified into two groups, those operating at the visual and photographic wavelengths and those at radio frequencies. Optical telescopes, including their design, operation, physical limitations and advantages, are the subject of this section.

Telescopes operating in the optical and photographic portions of the spectrum have two basic designs, the reflecting type, in which the main element is a mirror, and the refracting type, in which the main element is a lens. Table I(S7) lists the sizes and locations of some of the largest telescopes presently in operation.

TABLE I(S7)

LARGE TELESCOPES PRESENTLY IN USE (REF. 1)

Aperture Diameter (Inches)	Location
40*	Yerkes, Wisconsin
36*	Lick, California
200	Mount Palomar, California
120	Lick, California
104	Crimean Observatory
100	Mount Wilson, California
84	Kitt Peak, Arizona
82	McDonald, Texas
79	Jena, Germany
75	St. Michel, France
60	Mount Wilson, California
39	Burakan, Armenia

* Refracting type telescopes.

The telescope performs three basic functions: it gathers light, it magnifies, and it resolves. The magnifying power of a visual telescope is the ratio of the apparent diameter of an object as seen with the telescope to that as seen with the unaided eye, and is equal to the ratio of the focal length of the objective lens to the focal length of the eyepiece. In a well designed instrument it is also equal to the ratio of the diameter of the objective lens (or mirror) to the diameter of the exit pupil. The aperture of the objective is usually stated in denoting the size of a telescope, and the ratio of aperture to focal length (focal ratio) is about 1 to 15 or more. Thus, a 12-inch telescope is likely to be less than 15 feet long.

The brightness of the image of a point source is directly proportional to the area of the objective lens, or the square of the objective lens diameter. Thus, a point object would appear 25 times brighter with a 200-inch telescope than with a 40-inch telescope.

The resolving power of telescopes is commonly defined as the minimum angular separation of two point light sources such that they can be distinguished as two separate sources. Neglecting atmospheric effects, the diffraction limited resolving power of a telescope is governed by the diameter of the objective. The image of a point source of radiation is spread by diffraction into an Airy disk pattern, the angular size of which depends on the diameter of the objective.

Two criteria are commonly employed for specifying the resolving power of telescopes in terms of the minimum angular separation of the Airy disk patterns. The first is the Rayleigh criterion and the second is the half intensity criterion based on the coincidence of the half intensity points of the central diffraction peak.

(a) The Rayleigh Criterion - The distribution of illumination in the Airy diffraction pattern for a circular aperture is shown in Figure 1 (S7). Eighty-four percent of the total light is concentrated in the central disk.

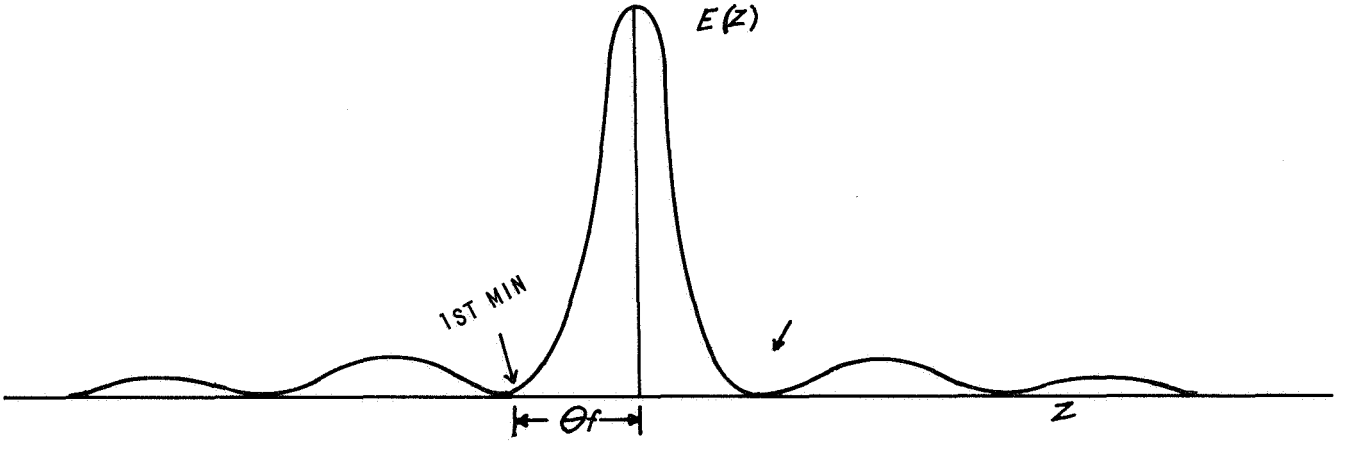


Figure 1 (S7) ILLUMINATION DISTRIBUTION AS A FUNCTION OF DISTANCE (REF. 2).

The maxima and minima may be represented as extremes of the diffraction intensity distribution of

$$I(z) = \left[\frac{J_1(z)}{z}\right]^2 \qquad \text{(S7-1)}$$

where $J_1(z)$ is the first order Bessel function. The Rayleigh criterion states that two point sources of equal intensity will just be resolved when the maximum of the central diffraction disk for one point source coincides with the first minimum of the diffraction disk for the other source. A solution of (S7-1) leads to the approximation for small angles, that the resolving power θ is given as

$$\theta = 1.22 \frac{\lambda}{d} \qquad \text{(S7-2)}$$

where d is the diameter of the lens. As an example, let $\lambda = 5000\text{Å}$, and $d = 10$ cm. Then

$$\theta = 1.38'' \qquad \text{(S7-3)}$$

The description of the image of a point source has been based on the tacit assumption that the light is monochromatic. For heterochromatic radiation the image consists of the superposed diffraction pattern produced by light of every wavelength. A salient feature not readily observable is that for heterochromatic light, the smallest central disk is violet but appears white to the eye, due to the combination of overlapping disks of light at all wavelengths[3].

(b) Half Light Intensity - Another criterion for resolution is based on half light intensity. When the images of two stars are closer together than the diameter of the central disk, they cannot be resolved as separate images. Here the diameter of the central disk is defined as the distance between points where the intensity of light is one-half that at the center of the disk [Figure 2(S7)]. The least distance in seconds of arc is given as

$$\phi = 1.03 \frac{\lambda}{d} \qquad \text{(S7-4)}$$

Using the same values for λ and d as in (S7-2), we obtain

$$\phi = 1.168'' \qquad \text{(S7-5)}$$

an 18% increase over (S7-3).

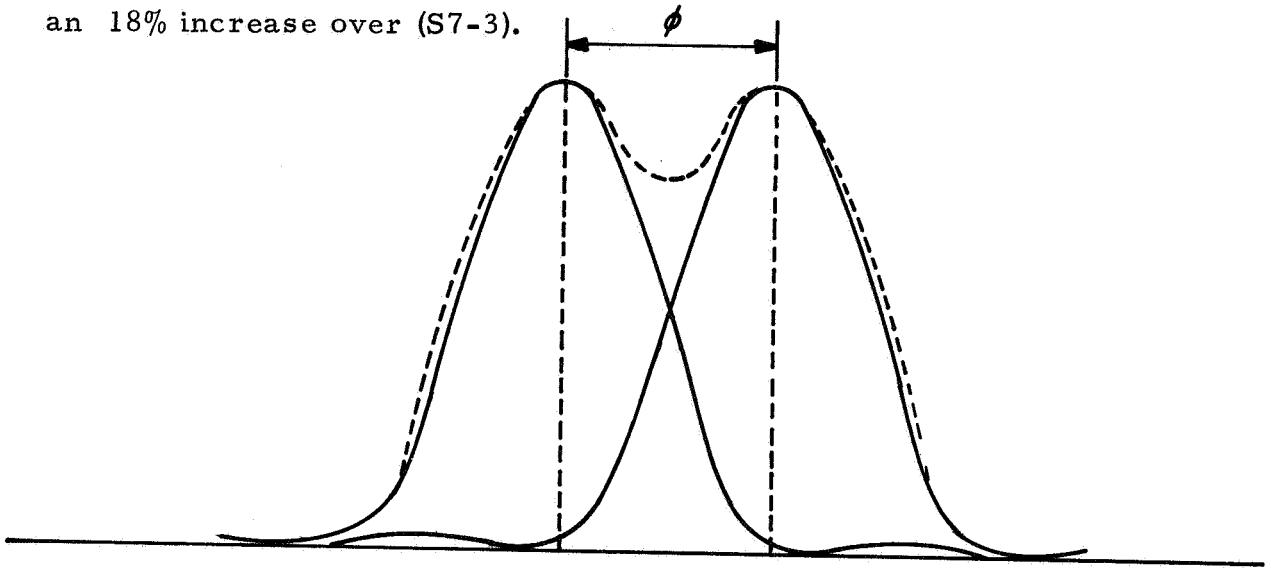


Figure 2 (S7) HALF LIGHT INTENSITY (REF. 2)

To illustrate the relationships of (S7-2), Table II(S7) is presented.

TABLE II(S7)

DIFFRACTION-LIMITED RESOLVING POWER
FOR SOME MAJOR TELESCOPES

Observatory	Aperture (Inches)	Resolving Power
Lick	36	0".127
Yerkes	40	0".114
Palomar	200	0".0228

The values obtained in Table II(S7) are for visual observations. When using photography, the least distances for minimum resolvable separation is made greater by the spreading of the images in the photographic emulsions.

A. Refracting Telescopes

The simple refracting telescopes[4], contain two double convex lenses separated by a distance equal to the sum of their focal lengths. The main element of this type telescope is the objective lens, which forms a real image. The linear size of the image of a particular object formed by an objective lens increases directly with the focal length of the objective lens. Its value is determined by the angular diameter of the object (in radians) times the focal length of the objective. Thus, an image formed by a 120-inch focal length telescope with an angular diameter of 0.01 radian will be about 1.2 inches.

B. Reflecting Telescopes

Five distinct advantages of mirror type telescopes over refractors have led to the incorporation of telescopes of a reflecting nature into recently constructed observatories. (1) The mirror is perfectly achromatic; there is no dispersion when light is reflected, (2) the mechanics of glass making become less strenuous. Striae and other defects in the disk, which would render it useless as a lens, do not make it unfit as a mirror, (3) the entire back of the mirror can be supported, whereas the lens is held only at its edge, (4) the optician has to figure only one surface instead of several, (5) the focal length of the mirror can be made shorter than that of the lens. The ratio of the aperture diameter to the focal length (focal ratio) becomes 1 to 5 or less. The primary mirror surface in reflecting telescopes is parabolic. The concave surface is coated with a thin film of metal, usually aluminum; hence the glass simply serves to give the required shape to the metal surface. Observations may be made at four different locations in the reflector type telescope[1]: Prime focus, Newtonian focus, Cassegrain focus, or the Coude' focus.

1. Prime Focus - The simplest reflecting type telescope is the Prime focus. The radiation beam is converged to its focal plane, where direct photographs are taken [Figure 3(S7)]. The 200-inch Hale telescope is of this type. The focal ratio is F3. 3, and since the diameter of the field of good definition of a paraboloid of this focal ratio is less than 0. 5 inch, a Ross correcting lens is normally introduced in front of the photographic plate to extend the field. By introducing a small telephoto effect into such a lens, it is possible to change the effective focal length of the system. Telephoto corrective lenses of focal ratio F3. 6 and F4. 7 have provided a well-corrected field about 3 inches in diameter for the 200-inch.

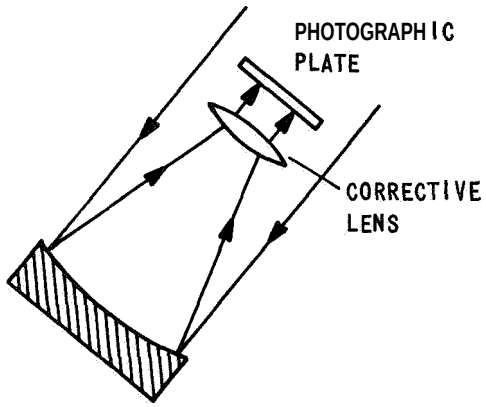


Figure 3 (S7) PRIME FOCUS WITH CORRECTIVE TELEPHOTO LENS

2. Cassegrainian Focus - In the Cassegranian type, the radiation reflected from the primary mirror is focused on a convex hyperboloidal mirror, which is located in front of the focal plane of the primary mirror, and the convergent beam is sent back through an opening in the center of the primary mirror [Figure 4(S7)]. The observer then looks in the direction of the source.

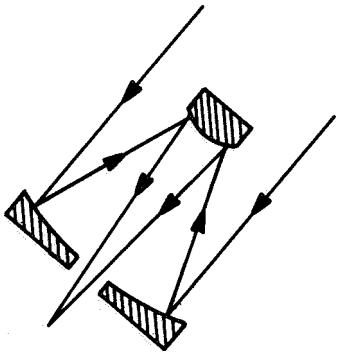


Figure 4 (S7) A CASSEGRAINIAN FOCUS TELESCOPE

3. Newtonian Focus - Replacing the hyperboloidal mirror by a plane flat mirror oriented such that the image is focused perpendicular to the longitudinal axis defines the Newtonian focus [Figure 5(S7)].

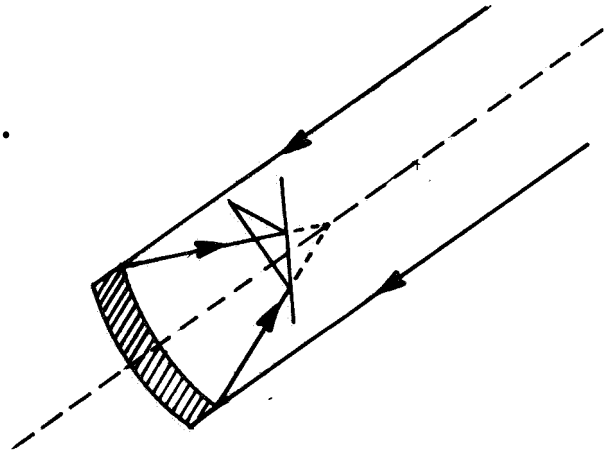


Figure 5 (S7) NEWTONIAN FOCUS

4. Coudk Focus - In the Coudé focus type, the radiation beam returned by the convex hyperboloidal mirror is diverted, by a small diagonal plane mirror, perpendicular to the longitudinal axis of the telescope [Figure 6(S7)].

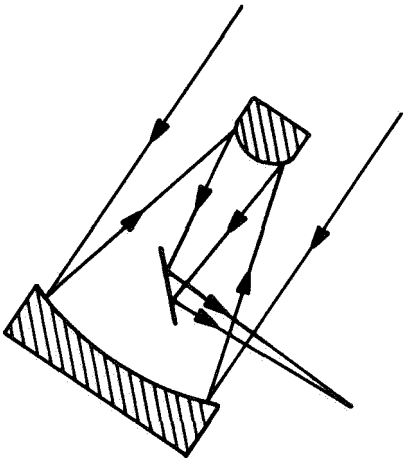


Figure 6 (S7) COUDÉ FOCUS

11. MIRROR MATERIALS

Materials used in the construction of reflecting mirrors must meet two basic criteria. The material must be hard enough to make it amenable to optical polishing and stable enough to allow optical precision to be obtained and maintained. [3]

Metal mirrors offer one outstanding advantage, namely, the large heat conductivity. In solar telescopes the conductivity becomes very important because of the absorption of some of the solar energy by the reflecting surfaces. For a metal film evaporated on glass, the conductivity is low, and the film will reach a temperature significantly above air temperature. This temperature difference will give rise to convection currents in the air that degrade the optical definition of the telescope. If the heat could be rapidly conducted through the mirror, the temperature rise could be kept small.

The basic requirement for hardness and stability eliminates many metals and alloys. Metallurgically, the single-element metals offer the best intrinsic stability, but most of them fail to meet the hardness requirement. Beryllium, titanium, aluminum alloys and kanogen appear promising.

A. Beryllium

This metal has several desirable physical characteristics, but is extremely toxic - more so than arsenic. It is light weight (ρ = 1.84 gm/cc) compared to Pyrexglass (ρ = 2.8 gm/cc) and very hard. Tests by the National Observatory show low granularity and small scattering. The extremely light weight of beryllium is attractive, where weight is of paramount importance.

B. Titanium

A pure metal of low density (ρ = 4.5 gm/cc) and high mechanical strength and hardness. Additionally, it has a low coefficient of expansion and high heat conductivity. Adversely, a polished surface of titanium shows a high degree of scattered light and is considered too porous to be used as a base metal when desiring a high-quality mirror surface, unless coated with other metals.

C. Aluminum Alloys

Both aluminum (ρ = 2.7 gm/cc) and copper (ρ = 8.9 gm/cc) are elements possessing high stability; however, they are too soft and ductile to be of practicable use.

Aluminum alloys, such as type 345, show ultrastable fine-grained characteristics. Anodizing the aluminum surface to Al_2O_3 yields an extremely hard oxide coating and a satisfactory mirror surface. The lightweight of mirror aluminum castings and their high heat conductivity and abrasion-free polished oxide faces offer a desirable combination. The chief disadvantage is the scattering caused by porosity in the anodized coating.

D. Kanogen

This material is an amorphous metal coating of nickel-nickel-phosphorous composition that at low temperature can be applied to a wide range of base materials. Three properties make it important for metal mirrors; it is amorphous, hard and chemically very stable. The hardness enables good polishing and the amorphous structure results in low scattering on a polished surface. The reflectivity is 50-60% so that low scattering usage is possible without coating the mirror with aluminum. Tests using a 0.003 inch thick kanogen coating on pure copper and aluminum disks show soft metal applicability.

III. EXO-ATMOSPHERIC ADVANTAGES

Observations from space stations located above the terrestrial atmosphere are advantageous for three reasons: they are free from the major part of infrared absorption by atmospheric water vapor and carbon dioxide; the angular resolving power with a telescope is not limited by atmospheric effects but rather by diffraction or aberration; and daylight observation of stars and planets is easier because of freedom from sky radiation, particularly in the red and infrared.

Freedom from infrared absorption permits determination of the amounts of water vapor and carbon dioxide in planetary atmospheres by the radiations they reflect to us. Also, infrared solar and stellar emission spectra may be studied, which could incorporate both the overall envelope of emission as it is influenced by hydrogen absorption and other effects, and details of spectra due to absorption by other molecules and atoms.

The atmosphere surrounding the earth imposes limitations on terrestrial based astronomical observations. The aperture effect* produces

* The illumination of the surface of an objective lens or mirror in the light of a single star is never completely uniform. The irregular patchwork of illumination which is actually present at any instant is called the shadow pattern. The rapid temporal changes in the shadow pattern are accompanied by changes in the integrated illumination, and scintillation results. The root-mean-square amplitude of scintillation tends to be larger for small aperture instruments. This is called the aperture effect.

image blurring and moving images. Meteorological phenomena, such as turbulence, prevent the attainment of the optimum theoretical resolving power and high field definition. A solution to the atmospheric problems may be obtained by placing the optical sensor at a sufficiently high altitude so that the path of the light through the air is a minimum. Even for larger aperture telescopes, 1" of arc resolution is a rare occurrence because of atmospheric effects and conditions which enable a 1" arc resolution are regarded as excellent.

An estimate based on work with a 6.5 inch refractor at Junipero Serra Peak Observatory indicates 1" seeing prevalent for 1% of the time, corresponding to four hours during a six week observation.

A. Balloon-Borne Exposure

M. Schwarzschild[5], using a 12-inch reflector with an effective focal length of 200 feet suspended from a balloon at an altitude of 80,000 ft., (96% above the atmosphere) obtained photographs of the solar granulation of unprecedented definition. Using exposure times of 0.001 sec, the theoretical optical resolution of 0.33 sec of arc was achieved over a field of 83" by 107" of arc.

B. Orbiting Astronomical Telescope

Studies have shown[6] that the requirements are within the present state of the art to orbit a 120-inch diffraction limited reflecting type telescope. The telescope is basically of Cassegrain configuration with a focal ratio of 8. The recommended primary mirror would have a beryllium metal base coated with Kanogen, with an estimated cost of $1,000,000. For images at the prime focus an angular field of 4 sec of arc and a linear field of 0.005 inch is predicted. Corrections to the mirrors indicate that all aberrations would fall within the Airy disk. The most critical alignment problem is that of keeping the secondary mirror on the optical axis. For a focal ratio of f/8, the centering tolerance is given as 0.0004 inch.

An analysis of the effect caused by therrnal gradients indicates that the mirror structure could be warped causing decentering of the secondary mirror which in turn causes astigmatism in both mirrors[6]. The maximum tolerable temperature difference between points on the opposite side of the structure was shown to be 8°F. Calculations indicate that the temperature gradient in the mirror will be less than that of the structure by a factor of about 200 steady state conditions.

Penetration of small micrometeoroid particles into the primary mirror would cause diffraction patterns which would be large compared to the Airy disk of the system. Assuming that the effect of solar radiation pressure exceeds the effect of solar gravitation in space, the micrometeoroids would be blown away from the solar system. Under this assumption, the life expectancy is calculated to be about 300 years[6].

The weight of the telescope, including supporting parts, electronics, alignment sensors, experiments and guidance control was found to be 20, 000 lbs.

For a 120-inch telescope, the angular resolving power (Rayleigh) is 0. 042" of arc at a wavelength of 0. 5 micron. This is the radius of the Airy disk to the first dark ring. In a Cassegrain system, the secondary mirror obscures part of the aperture; a 25% obscuration improves the resolving power by about 10%, i. e., the angular resolving power is 0. 03825" [Figure 7(S7)]. Additionally, central obscuration reduces the light in the central disk and increases the light in the rings around it, hence causing a loss in contrast. Without obscuration, the first bright ring is 1/57 as bright as the center of the disk; with 25% obscuration, it is 1/24 as bright. A method used to reduce the problem of obscuration is called apodization. In apodizing a system, the size and shape of the exit pupil is changed **or** the transmission amplitude of the outgoing wave is varied **or** the phase of the wave is locally changed. Apodization partially suppresses the amplitude height of the secondary maximum thus increasing contrast and resolution[7].

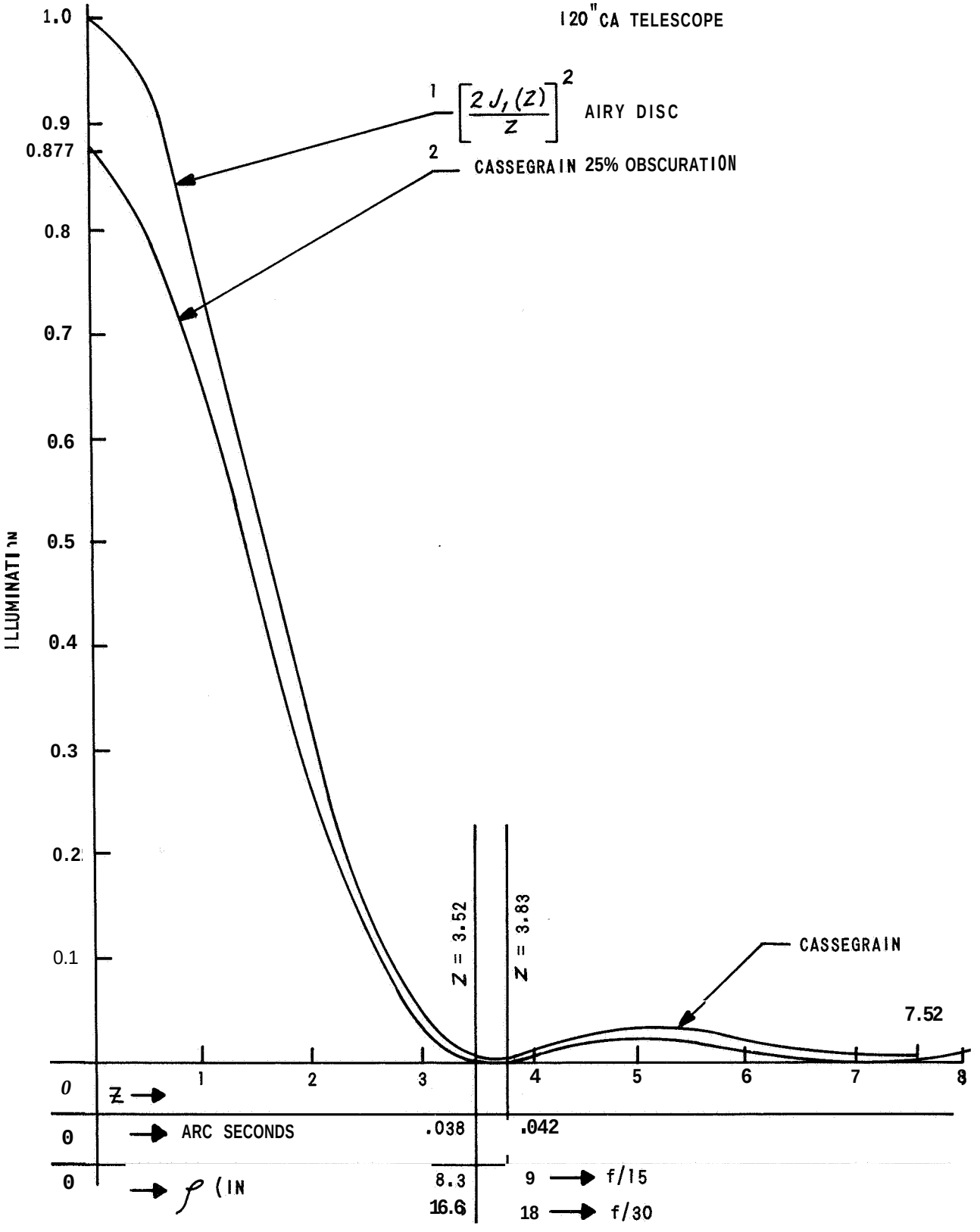


Figure 7 (S7) OBSCURATION EFFECTS ON A CASSEGRAINIAN TYPE TELESCOPE (REF. 6)

The amount of obscuration is dictated by the focal length of the primary mirror. In particular, high resolving power is apt to demand a long focal length to avoid the grain of the film. A large focal length would require large magnification at the secondary, approximately 20 to 50 times. Such a secondary mirror need not obscure more than 1/20 to 1/50 of the aperture.

IV. PHOTOGRAPHIC TECHNIQUES

The previous discussion on telescopes has had its accent on visual observations. For permanent recording or astronomical events, a variety of photographic techniques have been enlisted. Those factors which govern photographic emulsions, optical imaging devices and geometrical optics design in a space environment, are large and extensive and will not be examined here. For a comprehensive discussion on these parameters the reader is referred to References 4 and 8. It should be noted that, in general, earth-based astronomical observations, using photographic aids require exposure times up to and including several hours. When considering solar surface phenomena, an exposure period of this magnitude becomes unnecessary. Production of optical imagery of dynamical surfaces from spaceborne vehicles requires further investigation (see Appendix S-6).

V. SPECTROSCOPY

Spectroscopy utilizes at least half of the available observing time of most large telescopes[9]. When the source is of significant intensity, such as the sun, the dispersive mechanism is a grating with a system of fine parallel grooves ruled on a reflecting glass or metal surface. The resolving power being directly proportional to the number of grooves effective in producing the image. The better gratings contain up to 800 grooves per millimeter[10]. The use of a special form of groove shape (blazed) permits the main part of the total light to be confined to a certain angular range, thus giving a more favorable distribution of energy. To distinguish light waves, whose wavelengths are close together, the principal maxima of the wavelength formed by the grating should be as narrow as possible. The resolving power, R, is defined by

$$R = \frac{h}{A\lambda} \qquad \text{(S7-6)}$$

where λ is the mean wavelength of two spectrum lines that can just be recognized as separate and $\Delta\lambda$ is the wavelength difference between them.

The resolving power of a diffraction grating may be expressed as;[11]

$$R = mN \qquad \text{(S7-7)}$$

where N is the number of grooves on the grating and m is the order. For $\lambda = 5000\text{Å}$, an 8000 groove grating gives a first order resolution of $A\,h = 0.625\text{Å}$.

In the plane-grating spectrograph, monochromatic slit images are produced and only a small region of the spectrum is sharply defined at any one time. These restrictions may be avoided by use of a concave grating. In solar spectroscopes the concave grating becomes the collimating, dispersing and image forming element. A variety of geometrical mountings have been developed which have the concave grating, the entrance slit, and the diffracted spectrum lying on a circle (Rowland circle) which has as its diameter the radius of curvature of the grating blank, and to which the grating is tangent [Figure 8(S7)]. A common value for the diameter is 21 feet.

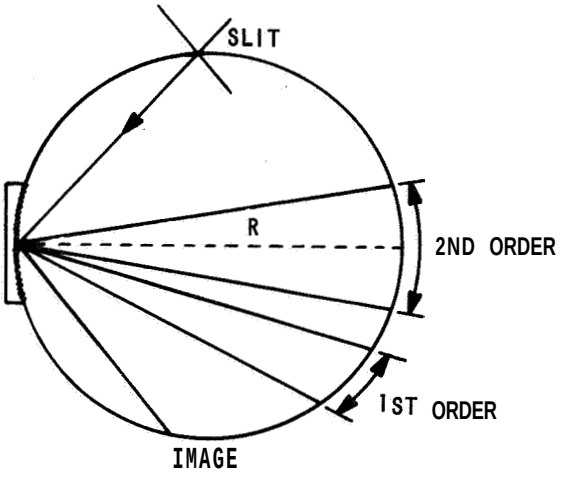


Figure 8 (S7) ROWLAND'S CIRCLE FOR CONCAVE GRATING (REF. 13)

Some advantages in this principle of construction for solar instruments are: (1)the number of optical elements is the absolute minimum one; (2) the system may be used for all wavelength regions; **(3)** spectra formed in the direction of the normal to the grating have constant and linear dispersion; **(4)** large ranges in wavelength and several spectral orders may be photographed simultaneously; (5) multiple slit and multiple gratings may be used to attain high efficiency in various regions. The principal disadvantage is the presence of a large degree of astigmatism in the image.

A variant of the concave grating spectroscope, the Wadsworth mounting[12], employs a concave mirror to collimate the radiation passing through the slit, thus achieving compactness and perfect achromatism. The slit is placed near the grating. The collimator and focal plane are close together at the other end of the instrument [Figure 9(S7)]. Astigmatism, coma, and spherical aberration are zero on the grating normal, and do not increase rapidly with distance from the normal. The perfection of the spectrum image makes it possible to correlate line widths and intensities in the spectrum point by point with the solar image on the slit. The spectra of small solar structures can be recorded and identified with this instrument.

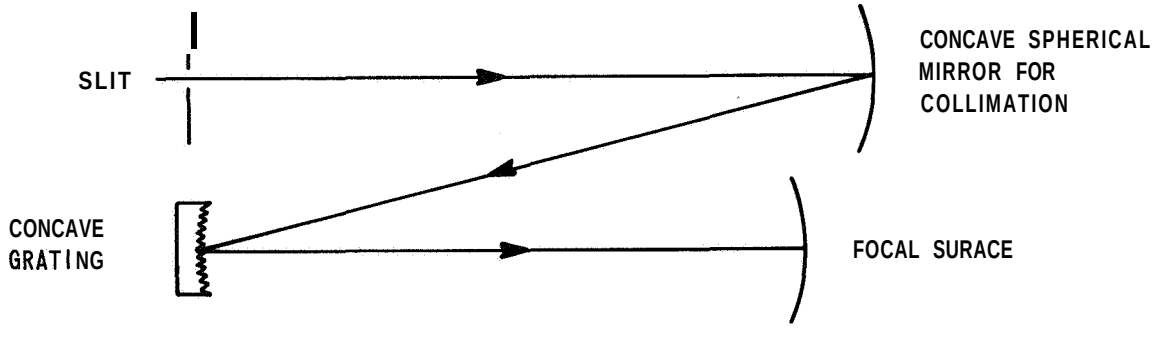


Figure 9 (S7) THE ALL-REFLECT1NG WARSWORTH MOUNT ING OF THE CONCAVE GRATING (REF. 12). THE FOCAL PLANE IS FOUND TO BE ABOUT HALF THE CIRCULAR DIAMETER

VI. SPECTROGRAPHS

An instrument used for photographing spectra is called a spectrograph. In astronomical applications the actual design of a spectrograph depends largely on the telescope with which it is used. A refractor has only one focus. In the case of a reflector, the spectrograph can be employed in various locations. Low-dispersion spectrographs are often attached to refractors. If used with a reflector, they are attached at the prime or the Newtonian focus. Spectrographs are effective in observing faint objects. Reflectors usually have f-numbers between f/3 and f/6 at the prime or Newtonian focus. Consequently, the collimator of a low dispersion spectrograph can be kept short, and the spectrograph as a whole is a fairly small and compact piece of equipment. For astronomical spectrographs of long focal length, folded collimators with mirror optics may be employed.

Gratings are usually mounted on a turret so that they can be rotated in order to image the desired order or region of the spectrum on the photographic plate. By a suitable choice of the grating properties, such as the number of grooves per millimeter and the blaze angle, the optical spectrum can be well covered with different dispersions.

The camera of a stellar spectrograph should be designed to yield the utmost speed. The proper choice of the camera for a spectrograph depends on many factors particularly on the linear dispersion desired. The linear dispersion increases with the focal length of the camera.

Modified microscope objectives with their small f-numbers have been used successfully for fast spectrograph cameras of low dispersion. Using an f-number 0.6, spectra of objects as faint as the 18^{th} magnitude have been made at $\lambda = 4350\text{Å}$, with a very low dispersion of 440Å/mm.

A consideration when adopting gratings for viewing a solar environment is their temperature sensitivity. Variations of temperature displace the spectrum lines owing to the change of grating space which results from the expansion or contraction of the grating. With a grating of speculum metal, a change of temperature of 0.1°C shifts the line of wavelength 5000Å in any order by 0.013Å [11].

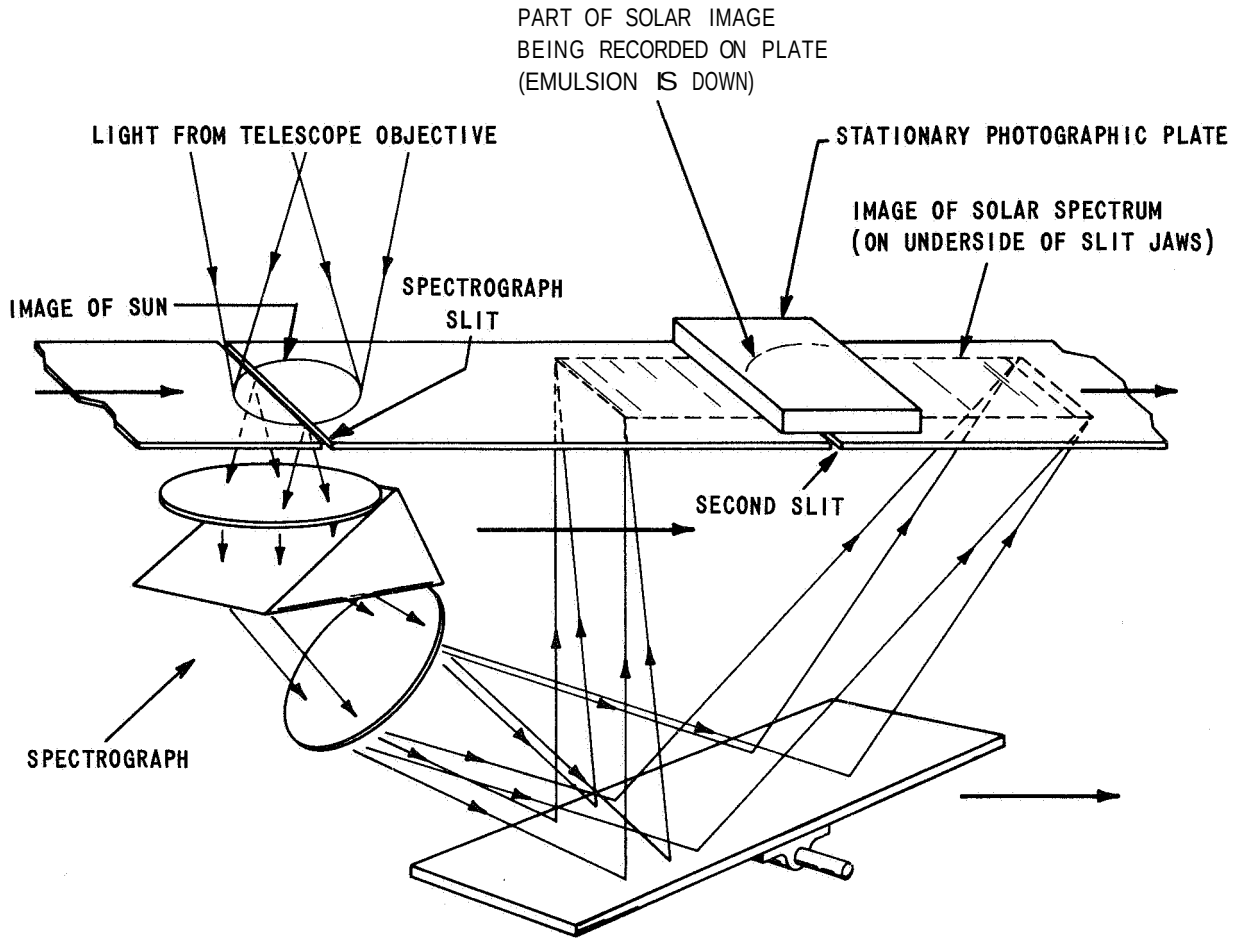
For investigations of the spectrum region below 2000Å, a vacuum spectrograph may be used. Enclosure of the spectrograph in a vacuum tank has the following advantages[12]; (1) all air currents within the instrument are eliminated; (2) the instrument is insensitive to temperature and pressure changes in the surrounding air; and (3) the mirrors and grating are protected from deterioration in the air.

To obtain a practical understanding of some spectrographic properties, consider the limiting characteristics of the Dunsink Observatory's spectrograph in Dublin, Ireland. A high resolving power of 1.1×10^5, or 70% of the theoretical one was achieved in the second order. Close pairs of Fraunhofer lines, such as $\lambda = 4355.9$Å and $\lambda = 4356.0$Å, $\lambda = 4401.45$Å and $\lambda = 4401.55$Å, or $\lambda = 4489.10$Å and $\lambda = 4489.19$Å are clearly resolved. The solar spectrum in this region, using the second order, required exposure times of 15 seconds, on slow Kodak B-4 plates[13].

VII. SPECTROHELIOGRAPHS

The modification of the spectroscope obtained by placing a slit in the focal surface so that the instrument transmits only radiation in a narrow band of wavelengths is called a monochromator. In this form, the spectroscope and a solar telescope combine to form a smooth continuous scanning motion, and an image of the Sun's disk, or a portion of it, in monochromatic light, can be built up in a photographic emulsion. By increasing the scanning rate, visual impressions of a monochromatic image may be obtained. An instrument with a rapid scanning frequency for visual use is called a spectrohelioscope [Figure 10(S7)]. For an extensive treatment of spectrohelioscopes and spectroheliographs, the reader is referred to Reference 14.

A recently completed (1962) spectroheliograph at the Crimean Astrophysical Observatory embodies a diffraction grating with a ruled surface 150 x 150 mm^2 with 600 grooves per mm. Monochromatic images of the Sun, 50 mm in diameter, can be produced simultaneously using radiation from any two spectral regions not closer together than 300Å. The passband of the instrument is generally one-tenth of an angstrom.



ENTIRE APPARATUS (EXCEPT PHOTOGRAPHIC PLATE) MOVES AS INDICATED. THUS, SPECTROGRAPH SLIT SCANS IMAGE OF SUN, AND SECOND SLIT, PASSING LIGHT IN DESIRED PART OF SPECTRUM, SCANS PLATE, SO THAT PHOTOGRAPH OF SUN IN THE LIGHT OF AN ISOLATED SPECTRAL REGION IS RECORDED.

MIRROR TO REFLECT SPECTRUM TO SECOND SLIT. ROTATION OF MIRROR CAUSES DIFFERENT PART OF SPECTRUM TO FALL THROUGH SECOND SLIT. THUS, DESIRED PART OF SPECTRUM CAN BE ISOLATED.

Figure 10 (S7) CONSTRUCTION OF A SPECTROHELIOGRAPH. (REF. 10)

A picture of the Sun made with a spectroheliograph, or any other monochromatic filter with a narrow passband, can record only those features whose line of sight velocities lie within a very narrow range, and rapidly changing objects cannot be observed often enough during their lifetimes in this way. Frequency suggestions in the literature have indicated that a zero dispersion monochromator would permit increasing the width of the passband filter to include a wider velocity range without degrading the quality of the image.

A multi-slit spectrograph, successfully used on planetary nebulae[15] appears applicable to prominence study, since they, like nebulae, radiate chiefly in a few widely separate emission lines. If one applies a polarizing monochromatic filter to project two monochromatic images in the plane normally occupied by the slit jaws of a spectroscope, and if one of the images is reversed from left to right with respect to the other, the two images will be re-imaged in the focal plane of the spectroscope with line-of-sight velocity shifts appearing to give stereoscopic relief to moving features[16].

A useful instrument in solar observations, which behaves like a filter with a very narrow transmission band, is the Fabry-Perot dtalon. It can be used to derive precise line contours[17] if one makes intensity measurements within the highly monochromatic channels obtained when the étalon is used with a spectroscope. Measurements with the étalon have achieved a precision of two or three parts in ten million for the average value of the wavelengths of solar lines. The same degree of accuracy is predicted for a grating and photoelectric recording combination[18].

VIII. PHOTOHELIOGRAPH

Telescopic cameras for taking pictures of the Sun are known as photoheliographs. An image resolution of 1" of arc or better is required for effective observations of the fine detail of the solar photosphere. One sec of arc corresponds to the theoretical limit of resolution of a 5-inch telescope.

One second of arc seeing implies conditions under which the photospheric granulation or any other fine detail can be distinguished both visually and on direct photographs taken with a short exposure time. Prolonged exposure for several seconds does not necessarily enable spectroscopic observation of equal resolution to be obtained because image degradation, shifts and distortions amounting to several seconds of arc are present. High resolution techniques using cinematographic arrangements have been recently developed. One such technique employs the 5-inch photoheliograph at Sydney, Australia. This consists of a 5-inch objective lens forming a

16 mm solar image at the prime focus. A second lens magnifies and produces an effective irnage diameter at the camera gate of 20 cm. An 800Å bandpass filter is centered on 5400Å. Exposure time is of the order of milliseconds.

IX. SPECIAL FILTERS

A polarizing monochromatic filter (birefringent) was designed by B. Lyot[19] and has been successfully used to study the solar corona and prominences without the necessity of a total solar-eclipse.

Lyot used six quartz plates varying from 2. 221 mm to 71. 080 mm in thickness with polarizers between each pair. The optic axes of all the plates were perpendicular to the light beam, and parallel to each other, while the polarizers were set at 45° to the optic axis. The filter transmitted 13 narrow bands with half-widths of 2Å. The undesired bands remaining were deleted by an ordinary colored filter. By varying the temperature of the filter, a shift in the wavelength of the transmitted bands may be accomplished.

X. MAGNETOGRAPHS

Spectral lines when produced in a magnetic field are split into distinct components (Zeeman effect). Two line components, circularly polarized, appear when the source is viewed in the longitudinal field direction and three line components, plane polarized, appear when the source is observed transversely to the field direction. The separation of the split components is directly proportional to the field strength at the location of the emitting or absorbing atoms. An instrument designed to measure magnetic field strength at the solar surface is called a solar magnetograph. When the direction of observation is parallel to the magnetic field, the two parts of the circularly polarized beam are polarized in opposite directions. It is these circularly polarized beams that are detected by the magnetograph, so that the magnetograph measures the component of the magnetic field that is parallel to the line-of-sight, in the following manner. The solar image is passed through a crystal which is

subjected to alternating voltage and alternately transmits right-circularly and left-circularly polarized light. The beam is passed through a single slit to a diffraction grating and dispersed with one of its lines centered on a double slit. The light passing through is converted into electric currents that vary the position of the trace on a recording oscilloscope. Thus the magnetograph separates the two components of a split line by viewing them alternately one at a time. The image of the line therefore oscillates back and forth over a small distance, the displacement depending on the strength of the magnetic field that produces the splitting [Figure 11(S7)].

As an example, consider the Mount Wilson solar magnetograph. The dispersion is 11 mm/Å and the corresponding Zeeman separation is a hundred thousandth of an angstrom, but each line in the unsplit spectrum is itself about a tenth of an angstrom wide. Thus, the parts into which a line is split almost completely overlap. The function of the magnetograph is to distinguish the overlapping lines and measure their separation.

The Mount Wilson magnetograph scans the whole of the Sun's disk once an hour with a resolution of 23 seconds of arc. It can measure magnetic fields as weak as 0. 3 gauss, although the usual working range is 1. 0 to 50.0 gauss or more.

Since the variation of intensity with wavelength is approximately linear for small distances along the contour of the edge of a spectral line, the amplitude of the modulated signal will vary linearly with field strength for small fields.

For large field values, between 1000 and 3000 gauss, like those found near sunspots, the shift is easily detectable. On the other hand, the shift when the magnetic field is 1. 0 gauss is 0. 00008Å, while the width of the line is 0. 1Å. In 1952, Horace W. and Harold D. Babcock, using a high dispersion spectrograph at the Hale Solar Laboratory in Pasadena, California developed equipment to measure photoelectrically shifts caused by fields smaller than 1. 0 gauss. Using the iron line, 5250. 2Å, they found that the equipment generated a noise level of the order of 0. 1 gauss, thus suppressing the minimum limit to measurements in the line shift of only 0. 0001 of its width[20].

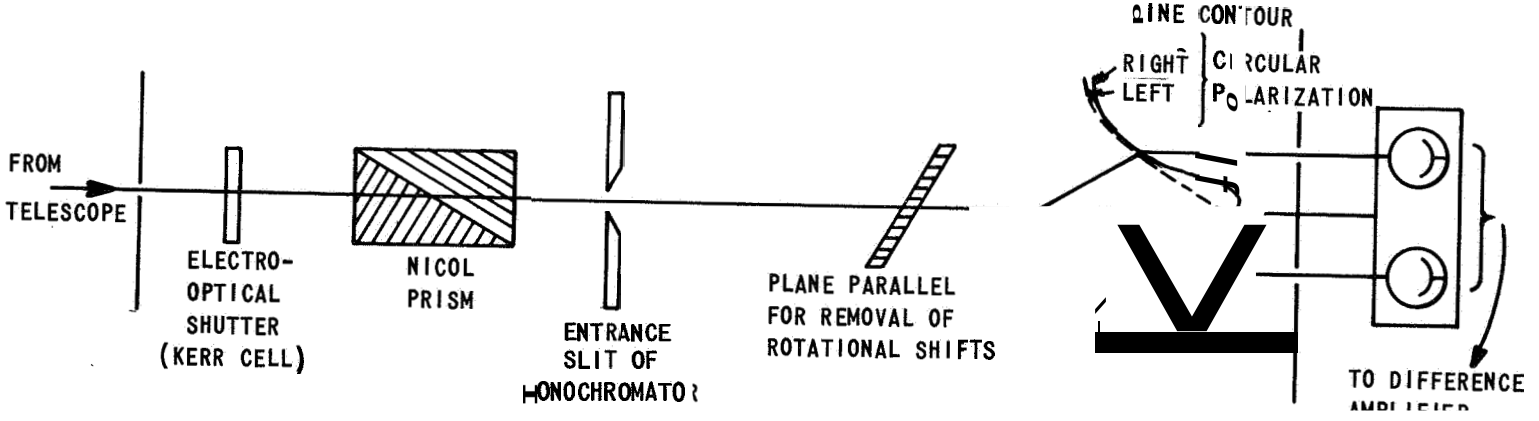


Figure 11 (S7) BABCOCK MAGNETOGRAPH AT MOUNT WILSON (REF. 18)

The investigators found that a general solar magnetic field, in latitudes within 20° of the solar pole, on the order of 1 gauss, existed. Additionally, regions in the sunspot belt had fields of the order of 3 gauss and were not associated with any visual features.

XI. CORONAGRAPH

The coronagraph is an instrument, designed by B. Lyot, to observe the corona of the Sun without an eclipse. A conical occulting disk is placed between the primary and achromatically corrected secondary lens of a refracting telescope. In this manner the Sun's disk is artifically eclipsed and spectroscopic study of bright corona emission lines may be achieved. Difficulty in the main lens construction has kept coronagraphs relatively small. The largest is at the Sacramento Peak Observatory and has a 16-inch aperture. The corona has strong emission lines in the green at $\lambda = 5303\text{Å}$ and in the red at $\lambda = 6374\text{Å}$ and $\lambda = 6702\text{Å}$. From a study of the spectrum, the temperature and electron density of the corona have been deduced. Photographs in white light have rarely been taken except at eclipses because of the high sky scattering. Photographs are predominately in the light of one of the emission lines. The coronagraph can be used with a slit and polarizing monochromatic filter called a polarimeter. A polarimeter produces a degree of discrimination against scattered sky light thus rendering measurements in the solar corona less difficult.

XII. DETECTORS

A multitude of instruments incorporate a photoelectric process in measuring astronomical radiation. The expansion of astrophysics, because of its dependence on the analysis of the radiation from astronomical bodies, has turned a large part of observational astronomy into an evaluation of intensities. The photoelectric method results in greater precision and has additional advantages because of the linear response and spectral range coverage by a single cathode.

The term physical photometer is applied to the family of radiation detectors in which the incident energy produces thermal or electrical effects that can be read directly on a scale or dial. Unlike the photographic plate, the response is usually directly proportional to the intensity.

Another group of detectors, all of which depend on the transformation of incident energy into heat, are non-selective and respond equally to radiant energy over a wide range of wavelengths. Examples of such detectors would include the pyrheliometer, bolometer, radiometer, radiation thermocouple and pneumatic cell. The range of wavelengths over which each of these is non-selective is determined by the transparency of the entrance windows and the spectral absorptivity of the receiving surface. For an extended treatment, the reader is referred to Reference 21.

XIII. CONCLUSIONS

Although many excellent instruments exist for solar observation and study from the surface of the earth, recent experimentation (higher resolution obtained with balloon-borne telescopes and ultraviolet spectra obtained with rocket-borne spectrometers) definitely indicates the desirability of making measurements of solar phenomena and parameters from above the atmosphere to avoid the degradations of absorption and turbulence. The complexity, size, and necessity of precise calibration (as indicated in this paper) suggest that instrument survival in a launch and space environment as well as proper functioning may be difficult to achieve. The costs of alternatives must be carefully weighed in light of the expected return before commitment to any particular instrument for solar observation from space can be made.

XIV. REFERENCES

1. Kuiper, G. P. and Middlehurst, B. M., Telescopes, University of Chicago Press, Chicago, 1960.

2. Ditchburn, R. W., Light, Interscience Publishers, Inc., New York, 1963, p. 278.

3. Hardy, A. and Perrin, F., The Principle of Optics, McGraw-Hill Book Co., Inc., New York, 1932, p. 129.

4. Baker, R., Astronomy, 8th Ed., Van Nostrand Co., New Jersey, 1964, p. 112.

5. Schwarzschild, M., Ap. J. , 130, 1959, p. 345.

6. American Optical Co., Feasibility Study of a 120-inch Orbiting Astronomical Telescope, 1963.

7. Born, M., Wolf, E., Principles of Optics, MacMillian Company, New York, 1964, p. 417.

8. Bsum, W. A. , The Detection and Measurement of Faint Astronomical Sources, Ed. Hiltner, W., Vol. II, "Stars and Stellar Systems, Compendium of Astronomy", University of Chicago Press, Chicago, 1962.

9. Abell, G., Exploration of the Universe, Holt, Rinehart and Winston, New York, 1964, p. 156.

10. Motz, Lloyd, Ed., Astronomy A to Z, Grosset and Dunlap, New York, 1964, p. 236.

11. Jenkins, F. A. , White, H. E. , Fundamentals of Optics, McGraw-Hill Book Co., New York, 1957, p. 349.

12. McMath, R. R., Mohler, O. C., Solar Instruments, Handbuch der Physik, Vol. LIV, Springer-Verlag, Berlin, 1962, p. 21.

13. Bruck, H.A. and Bruck, M. T., Vistas in Astronomy, Vol. I, Ed. by A. Beer, Pergamon Press, London, p. 437.

14. Wayman, P.A., Vistas in Astronomy, Vol. I, Ed. by A. Beer, Pergamon Press, London and New York, 1955, p. 422.

15. Wilson, O.C., Science, Washington, D.C., 122, 1955, p. 882.

16. McMath and Mohler, p. 28.

17. Shane, C.D., Bull. Lock Obs., 16, 1932, p. 76.

18. McMath and Mohler, p. 31.

19. Lyot, B., Ann d'Astrophys., Vol. 7, 1944, p. 1.

20. Miczaika, G. R. and Sinton, W. M., Tools of the Astronomer, Harvard University Press, Cambridge, Mass., 1961, p. 257.

21. Smith, R.A., Jones, F. E. and Chasmar, R. P., The Detection and Measurement of Infrared Radiation, Oxford, 1957.

APPENDIX S-8

AN ACTIVE PROBE OF THE SUN

By David A. Richards

This appendix presents the results of a study and analysis of a laser application to the investigation of the sun in an active probe mode. It was hypothesized that a high power laser might be utilized to obtain information about the physical parameters measurable in the solar atmosphere by analysis of the reflected light. The mechanisms by which light could be reflected (backscattered) were examined to obtain quantitative comparison with the magnitude of the background noise. Calculations using the best available values for the system constraints found such a system wanting by several orders of magnitude. It was concluded that a laser probe was not appropriate for such application. However, it was found that microwaves could probably produce useful information about the corona when used as an active probe.

I INTRODUCTION

There exists the possibility of obtaining information about a volume of gas by recording and analyzing optical energy scattered from it (see Ref. 1). This task is relatively simpler when the source of illumination is under the experimenter's control. With the advent of high energy pulsed lasers many applications to such problems are being advanced. One such application that has been suggested is the use of a laser in an active probe of the solar atmosphere. Although the background noise (fluctuations of the solar emission) immediately appears as a great problem such an approach cannot be dismissed until numerical values are considered. The following sections outline a tentative system, describe the mechanisms considered, present the results of calculations for signal and noise levels and discuss the possibility of using microwaves for an active probe.

II. SYSTEM CHOICE

The analysis presented here is intended only to give orders of magnitude answers to determine feasibility so that the range of values obtainable for the system parameters and constraints were researched only to that extent. The basic concept utilized for this study involves the use of a giant pulsed laser at a station close to the sun to illuminate a relatively small beam path through the solar atmosphere. By the analysis of the time and intensity (above background) of return, an attempt to learn something about the size, density and nature of the particles in the path can be made. Such an arrangement is similar to the laser rangefinder considered for use as an altimeter. The main differences are the continuous character of the return and the larger background level. The probe station was chosen at ten solar radii (~0.05 astronomical units) from the "surface". This choice is based on the capability of a vehicle to survive in the close proximity of the sun and is discussed further in Appendix S-5.

To obtain range resolution a choice of pulse length must be made. For electromagnetic radiation considered to be travelling at the speed of light in vacuo (3×10^8 m/sec) a microsecond pulse has ~150 meter resolution in range. Two hundred million watts average power in pulses with a duration of 30 nanoseconds have been achieved (see Ref. 2) using a Q-switched ruby laser. Two thousand joule pulses are quoted as the currently available maximum in Reference 3. It was decided that one kilometer range resolution (~6.7 μ sec) using a two thousand joule pulse would constitute an adequate estimate of laser pulse state of the art for this application.

The beamwidth or beam divergence of a pulsed laser is considered good at ~10^{-3} radians. To allow for improvements in this and other parameters a value of one second of arc (~5×10^{-6} radians) was chosen for this hypothetical system. Such a beam would illuminate a spot on the photosphere of approximately 10^9 meter^2 at the distance quoted above (7.5×10^9 meters).

A one angstrom bandwidth centered on the laser frequency was chosen for the receiver to accommodate the Doppler broadening due to scattering (~1/5Å). At present, pulsed lasers operate from 2700 angstroms to 10.6 microns. The ruby laser wavelength at 6943 angstroms was considered representative primarily because of the great deal of **work** that has been done at this wavelength for generating giant pulses.

III. REFLECTING MECHANISMS

Three primary mechanisms were considered as possible sources of reflection, namely, absorption and re-emission, Thomson scattering by free electrons, and Rayleigh scattering by the atoms.

The primary absorber present in the solar atmosphere is considered to be the negative ion of hydrogen. At the relatively low temperatures (compared to the center) of the atmosphere, neutral hydrogen atoms acquire another electron available from ionized metals, The photoionization of these ions is a source of continuous photon absorption. This is bound-free absorption. Free-free absorption in the field of the neutral atom is also possible. The combined continuous absorption curve as derived from quantum mechanical calculations is presented in Appendix **S-4. For** a beam of monochromatic radiation, absorption not only disperses the re-emitted radiation into 4π steradians but also re-emits it in a wide band of wavelengths.

The free electrons present from the ionization of the metals can give rise to Thomson scattering of incident radiation. The electron follows the oscillation of the applied field and radiates as a dipole at the same frequency. This radiation is not isotropic but goes as $\frac{1}{2}(1 + \cos^2\theta)$ where θ is the scattering angle. The total cross section is a constant and not dependent on wavelength.

For the wavelength considered above, Rayleigh scattering can occur from the atomic species present. This will be primarily hydrogen but helium was also considered. Rayleigh scattering applies to particles which are small compared to the wavelength. The cross section goes as the inverse fourth power of the wavelength.

IV. CALCULATIONS

To obtain the range resolution mentioned in Section II requires the integration time of the receiver not to be greater than the length of the pulse. The return will be continuous because of the translucence of the medium involved. Several simplifying assumptions have been made in favor of the system to facilitate the calculations of the orders of magnitude for signal and noise. No attenuation was considered to affect the return. All the energy extracted from the pulse was considered re-radiated by reflection.

The magnitude of the background noise was taken as -8600 watts/m^2/A" at 6943Å at the surface (zero level) of the photosphere. This is a conservative approximation based on solar data.

The Thomson scattering cross section is used for the scattering from electrons,

$$\sigma_T = 0.66 \times 10^{-24} cm^2 \qquad \text{(S8-1)}$$

An average value of the absorption cross section for the negative hydrogen ion was taken as

$$\sigma_A = 2.6 \times 10^{-17} cm^2 \qquad \text{(S8-2)}$$

The Rayleigh scattering cross section applicable to the photosphere was derived as follows:

$$\sigma_R = \frac{8}{3} \pi k^4 \alpha^2 \text{ where } k = \frac{2\pi}{\lambda} \qquad \text{(S8-3)}$$

where α is the polarizability of the particles and λ is the wavelength. α can be derived from

$$n = 1 + 2\pi\alpha N \qquad \text{(S8-4)}$$

where n is the index of refraction and N is the particle density. Using $n = 1 + 10^{-6}$ as given in Reference 4 and $N = 10^{17}\,\text{cm}^{-3}$ as given in Reference 5, we have

$$\alpha = 1.6 \times 10^{-24}\,\text{cm}^3 \qquad \text{(S8-5)}$$

which closely approximates the polarizability of monoatomic hydrogen. From this and Equation **(S8-3)** we have

$$\sigma_i = 1.4 \times 10^{-27}\,\text{cm}^2 \qquad \textbf{(S8-}6\textit{)}$$

The energy extracted from the pulse as it traverses a distance Δx can now be computed. If the irradiance of an area is I_o and the irradiance of an area **A %** deeper is I

$$\Delta I = I_o - I \simeq I_o\, \Delta x \sum_i \sigma_i N_i \qquad \text{(S8-7)}$$

where σ_i is the cross section of the i^{th} mechanism and N_i is the particle density of the i^{th} particle. The number of electrons in this region are given as

$$N_e = 10^{13}\,\text{cm}^{-3} \qquad \text{(Ref. 5)} \qquad \text{(S8-8)}$$

The negative hydrogen ion concentration is given as

$$N_{H^-} = 10^{9}\,\text{cm}^{-3} \qquad \text{(Ref. 6)} \qquad \text{(S8-9)}$$

The Rayleigh scatterers have already been given as

$$N_s = 10^{17} cm^{-3} \tag{S8-10}$$

With I_o = 300 watts/m^2 and $A\alpha$ = 1 km, equation (S8-7) gives

$$\Delta I \simeq 1 watt/m^2 \tag{S8-11}$$

If we considered the layer a partial mirror the fraction reflected in 1/300.

If we are to use this return for measurement, it is now appropriate to consider the signal-to-noise ratio achievable. The signal-to-noise ratio is generally given on a power basis. The signal power measured will be proportional to the square of the photoelectron current produced by signal photons. The noise will be due to the quantum fluctuations in the background photon stream, and the power associated with this is proportional to the mean square of the fluctuation which is simply the mean for photoelectrons described by a Poisson distribution (see Ref. 7). We then have

$$\frac{S}{N} = \frac{n_s^2}{n_B} \tag{S8-12}$$

where n_s is the number of signal-produced photoelectrons, and n_B is the number of background-produced photoelectrons during an equivalent time. One can derive both these quantities according to the following equations,

$$n_s = \frac{\eta T \lambda P_s \tau}{hc} \tag{S8-13}$$

$$P_s = (\Delta I) \frac{A_s}{2\pi} \frac{A}{R^2} \tag{S8-14}$$

$$n_B = \frac{\eta T \lambda P_B \tau}{h c} \qquad \text{(S8-15)}$$

$$P_B = I_B^{\lambda} \frac{A_s}{I n} A h \frac{A}{R^2} \qquad \text{(S8-16)}$$

where η is the quantum efficiency of the photoconverter,

T is the combined transmission factor of optics, filter, etc.,

λ is the wavelength of the laser,

P_s is the received signal power,

ΔI is the reflected flux from the photosphere,

A_s is the area of'the beam cross-section at the photosphere,

A is the receiver aperture,

R is the distance from photosphere to receiver,

P_B is the received background power,

τ is the integration time of the receiver,

h is Planck's constant, 6. 625 x 10^{-34} joule sec. ,

c is the speed of light, 3 x 10^8 m/sec,

I_B^{λ} is the solar spectral flux at the surface, and

$\Delta\lambda$ is the passband of the receiver centered on λ .

Using the following values for these parameters,

$7 = 0.05$, $\lambda = 0.6943\ \mu$

$T = 0.5$ $\Delta I = 1$ watt/m^2

$A_s = 10^9 m^2$ $R = 7.5 \times 10^9 m$

$\tau = 6.7\ \mu$sec $I_B^{\lambda} \simeq 8600$ watts/m^2"A"

$A h = 1\text{Å}$

one obtains a signal-to-noise ratio of

$$S/N \simeq 2 \times 10^{-4} A \quad \text{where} \quad A \quad \text{is in meters}^2 \qquad (S8\text{-}17)$$

The 200 inch reflector of Mt. Palomar has an aperture area of approximately $20 m^2$. Using this, and remembering that our return signal calculation was overly generous, we see that our desired signal would be buried in background noise.

Fluctuations of the background due to temperature differences across the granulation and the great difficulty of pointing accurately to receive return only from the illuminated field add to the conclusion that such a scheme is unfeasible.

V. MICROWAVES

From the above considerations it appears that we will have to be satisfied with peripheral attack upon the sun with active probes as a frontal assault is overwhelmed by the enormous output of energy in the visible region of the electromagnetic spectrum. However, the microwave region offers more hope. The peculiar characteristic of charged particle distributions known as the plasma frequency, results in electromagnetic radiations of higher frequency being totally reflected. This mechanism leads to a relatively high return signal and denies emanations from below to be transmitted outward so that detectable noise occurs only above the reflection surface. The electron density distribution of the corona would .be a prime candidate for study by this means. Calculations of the signal-to-noise ratio for radar reflection from the sun for an earth-based system were made by Kerr in 1952 (see Ref. 8). He concluded that "a system of this kind is within the range of present techniques, but it is obvious that a project of considerable magnitude is involved". At the much closer distance a probe could achieve, new calculations of a similar nature would be required. Since the received power goes as the inverse fourth power of the distance an enormous gain can be anticipated in this area. The antenna size, frequencies available, and transmitter power as well as' many other parameters would also require consideration.

VI. CONCLUSIONS

It is concluded that a backscatter laser probe of the sun is not feasible as a means of obtaining information about the solar atmosphere because of the enormous power requirement to achieve a satisfactory signal-to-noise ratio for the return. It is cautioned, however, that this analysis did not consider all possible applications of an active laser probe to solar research. Other modes of operation such as transmission attenuation through the corona require their own analysis in respect to feasibility and value.

For the backscatter probe it was also concluded that microwaves offer far greater hope of success than visible light. The information obtainable in this manner would be of value and such a system should be examined in greater detail.

VII. REFERENCES

1. Van de Hulst, H. C., "Light Scattering by Small Particles", John Wiley & Sons, Inc., New York, 1957.

2. Laser Focus, October 1966, p. 17.

3. Laser Focus Supplement, Characteristics of Lasers, 1965.

4. Kuiper, G. P., Ed., "The Sun", The University of Chicago Press, Chicago, 1953, p. 130.

5. Pawsey, J. L. and R. N. Bracewell, "Radio Astronomy", Oxford University Press, Oxford, 1955, p. 110.

6. Aller, L. H., "Astrophysics, The Atmosphere of the Sun and Stars", Ronald Press Co., New York, 1963.

7. Oliver, B. M., "Thermal and Quantum Noise", Proc. of IEEE, 53, May 1965, pp. 436-454.

8. Kerr, F. J., "On the Possibility of Obtaining Radar Echoes from the Sun and Planets", Proc. of the IRE, June 1952, Vol. 40, pp. 660-666.

APPENDIX S-9

FOURIER TRANSFORM SPECTROMETRY WITH RESPECT TO SPACE VEHICLE BORNE TELESCOPES

By John D. Gallatin

SUMMARY

This appendix serves as an introduction to Fourier transform spectrometry and discusses some of the problems of implementing it on a space vehicle.

I. INTRODUCTION

Fourier transform spectrometry (FTS) is a method of obtaining a spectrogram by harmonically analyzing the output of an interferometer. It differs from Michelson's work in that the analysis is performed on the on-axis irradiance rather than on the contrast of the fringes.

Section II below shows how it is possible to obtain spectra by use of an interferometer and how the resolution and spectral range are related to the interferometer parameters and the method in which it is used. Sections III and IV below show how, in some situations, FTS can provide spectrograms with less noise, ease stabilization requirements, or derive data which may be more helpful in testing theories and models. Comparisons with grating spectrometry are included to show in which cases FTS would be most advantageous. Comparisons of gratings and prisms are available in standard texts.[1] Section V contains a discussion of problems of implementation of FTS, especially in space. Sections VI and VII contain conclusions and recommendations based on the above discussions.

11. THEORY OF FTS

A. Relationship Between Spectra and Interferometer Records

Within scalar field theory[2], a spectrogram may be considered as a constant times the power spectrum of the electric field over some range of frequencies. The classical method of measuring spectrograms is to separate the frequencies and measure the power of each one. In that method, the grating or prism acts as a filter.

The FTS method is an alternate way of obtaining the power spectrum and uses the fact that the power spectrum is the Fourier transform of a constant times the autocorrelation function of the electric field.[3] The autocorrelation function may be determined by suitable mathematical processing of the output of a suitable interferometer and the spectrum obtained by harmonic analysis of the autocorrelation function.

Consider an electric field $\mathcal{E}_o(t)$ passing through a Michelson interferometer (see Figure 1) having a symmetrical beam splitter and two identical mirrors. Then the electric field at the output is the sum of that arising from each path through the interferometer. Considering the path to mirror #1, the output field is

$$\mathcal{E}_1(t) = \overleftarrow{r_B}(t) * r_1(t) * \overrightarrow{t_B}(t)\, \mathcal{E}_o\left(t - \frac{d_1}{c}\right) \qquad \text{(S9-1)}$$

where

$*$	=	denotes convolution.
$\overrightarrow{t_B}(t)$	=	transmittive impulse response of beam splitter for radiation incident from the left.
$r_1(t)$	=	reflective impulse response of mirror #1.
$\overleftarrow{r_B}(t)$	=	reflective impulse response of beam splitter for radiation incident from the right.
d_1	=	group optical length of path #1, i.e., optical path based on group velocity.
c	=	vacuum speed of light.

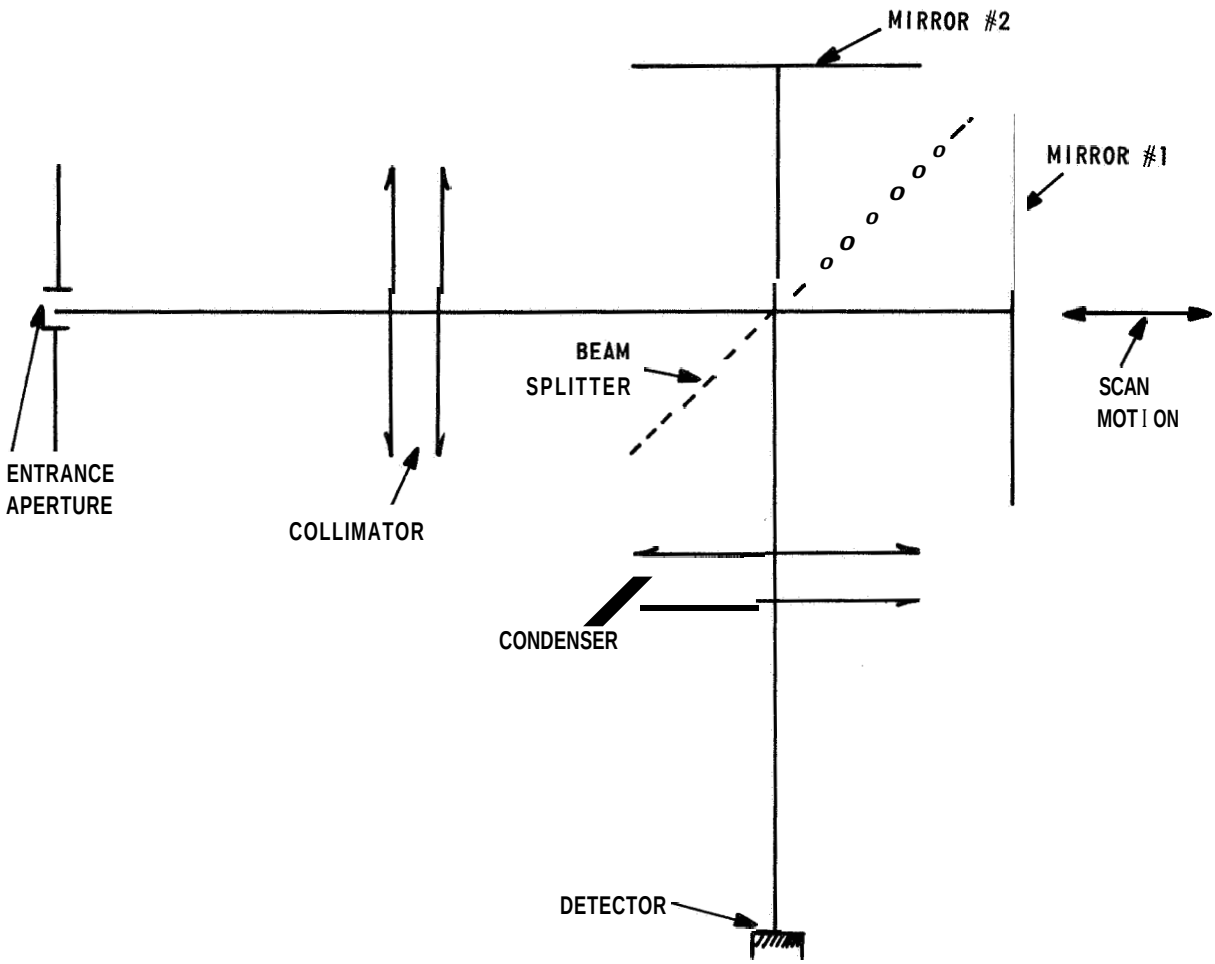


Figure 1 (S9) MICHELSON INTERFEROMETER

It is necessary to consider the possibility that the beam splitter and mirrors have impulse responses other than a scaled delta-function in order to arrive at meaningful results when one considers operation over a broad spectral range. If the beam splitter is symmetrical, and the mirrors identical as assumed here, the only effect of the impulse response is that it is equivalent to a color filter. When this is not the case, the analysis is considerably more complicated. For path #2

$$\mathcal{E}_2(t) = \vec{t_B}(t) * r_2(t) * \vec{r_B}(t)\, \mathcal{E}_o\left(t - \frac{d_2}{c}\right) \qquad \text{(S9-2)}$$

Because it was assumed that the mirrors are identical and the beam splitter is symmetrical, the total field is

$$\mathcal{E}_{\Sigma}(t) = t_B(t) * r_M(t) * r_B(t) \left[\mathcal{E}_o\left(t - \frac{d_1}{c}\right) + \mathcal{E}_o\left(t - \frac{d_2}{c}\right)\right] \qquad \text{(S9-3)}$$

where simplifications in notation have been made.

Radiation detectors at the frequencies of interest respond to average power, not field. Thus, the detector response will be proportional to the time averaged square of the above quantity.

Inspection of the time average of the square of the above quantity shows that it is composed of two parts. The first part is proportional to the input power, and the second is proportional to the autocorrelation value corresponding to $(d_1 - d_2)/c$. Hence, the required autocorrelation function may be found by scanning one of the mirrors, recording the irradiance and subtracting the proper correction.

B. Spectral Range of FTS

When records of interferometer irradiance as a function of path difference - these are called interferograms - are made and used to compute spectrograms having both high spectral resolution and broad spectral range, i. e., there are many independent wavelength resolution invervals, it is customary to perform the harmonic analysis by use of a digital computer. This avoids difficulties in fabricating analog devices with the required time-bandwidth product and adequately low noise. Because such processing and certain types of telemetry use only the interferogram values at a finite set of points, usually equally spaced, say with an increment Δx of mirror movement, the phenomenon of aliasing[3], also known as folding, must be considered. Folding requires that an interferometer be supplied with filters (or detector cutoff) to limit the detected radiation to a fixed spectral range, and is analogous to the requirement that a grating be supplied with radiation covering only one order.

The mechanism of folding is the same here as discussed in Ref. 3, but an independent derivation is provided to avoid confusion over certain numerical factors.

The on-axis irradiance for a single wavelength is proportional to

$$I \propto \cos \frac{4\pi x}{\lambda} \tag{S9-4}$$

where

x = displacement of one mirror from image of other

λ = wavelength

and it is assumed that the interferometer is in air or vacuum. However, the quantity $\left(1 + \cos \frac{4\pi x}{\lambda}\right)$ is equal to

$$\left[1 + \cos\left(2\pi M \pm \frac{4\pi x}{\lambda}\right)\right]$$

where M = any integer.

Expressing this quantity in terms of wave number, ν, (the reciprocal of wavelength) and taking the separation x as an integer j times the separation increment Δx we have

$$1 + \cos(4\pi \nu j \Delta x) = 1 + \cos(2\pi M \pm 4\pi \nu j \Delta x) \tag{S9-5}$$

If K is an integer, then $M = Kj$ will be an integer for all j. Thus for all j, $4\pi j \Delta x$ may be confused with $2\pi j K \pm 4\pi j \nu \Delta x$ or ν maybe confusedwith $\frac{K}{2\Delta x} \pm \nu$. Thus, to secure unambiguous results, ν must be restricted to one of the ranges 0 to $\frac{1}{4\Delta x}$ or $\frac{1}{4K}$ to $\frac{1}{2\Delta x}$, etc. The ability of radiation at one wave number to appear at another is the reason for the name aliasing.

C. Resolution of Spectra

The spectral resolution is a measure of the width of the spectral window, which is also called the instrument profile. All real instruments measuring real sources will blur the spectrum, and the blurring is equivalent to convolving the spectrum with the spectral window. It may be measured by attempting to measure the spectrum of a source whose radiation is confined to a band very narrow compared to the spectral window, and translating the apparent spectrum to zero wavenumber.

If FTS is performed by using equal weighting of all interferogram ordinates (after subtraction of DC) prior to the harmonic analysis, then the profile will be (assuming collimated light)

$$P_1(\nu) = \frac{\sin 4\pi \bar{x} \nu}{4\pi \bar{x} \nu} \qquad (S9\text{-}6)$$

where

ν = wave number

$\bar{x}$ = maximum separation of one mirror from image of other in interferometer.

The functional form of $P(\nu)$ is a diffraction function, as expected because it is the transform of a square pulse lag window.[3] The scale may be found by finding the first zero of $P_1(\nu)$ by using the orthogonality relation in cosine transforms.

Use of the above profile is often dangerous because of the large side lobes. Other profiles may be obtained by apodization. For example, a profile which is a diffraction function squared, rnay be obtained by multiplying all of the interferogram ordinates (prior to harmonic analysis but after subtraction of DC) by a linear function of x which has the value of unity for $x = 0$ and equals 0 at $x = \bar{x}$. The diffraction function squared profile has lower side lobes, but either $\bar{x}$ or the value of ν for which $P(\nu)$ first equals zero will be doubled. It is to be noted that the above profile is that resulting from a finite maximum mirror separation, assuming collimated light. Usually the spectrum will be further blurred by use of a finite source size.

III. APERTURE ADVANTAGE AND DETECTOR NOISE

In this and the following section, FTS is compared to grating spectrometry on the basis of signal-to-noise ratio. Depending on the measurements to be made, there are *two* reasons for expecting possibly better signal-to-noise ratio by use of FTS. First, it is sometimes possible to get more radiation into the interferometer. Second, in some cases, the interferometer method is more efficient. The first of these considerations is discussed in this section, the second is deferred to the next section.

For an extended source and radiation covering a narrow wavelength band $\Delta\lambda$, the power accepted by an optical instrument is given by

$$P = N_\lambda \Delta\lambda A \Omega \cos\alpha \qquad \text{(S9-7)}$$

where

N_λ = spectral radiance, a characteristics of the source

A = area of one of the limiting apertures

α = angle between normal to this aperture and beam

Ω = solid angle subtended by the other limiting aperture, or its image as seen from the first mentioned limiting aperture.

A comparison of the product $A\Omega\cos\alpha$ which is sometimes called luminosity or "throughput" is discussed in this section. A is taken as the area of the grating or beam splitter, and Ω is the solid angle subtended by the slit or other entrance aperture. For both spectrometers and interferometers, the area A is limited by mechanical problems; fabrication, installation, and accurate motion. The product of solid angle and cosine of off-axis angle $\Omega\cos\theta$ must be limited to avoid compromising the spectral resolution. The accurate motion required of a grating is rotation about an axis parallel to the slits through accurately measured angles. The accurate motion required in the interferometer is linear motion of one of the mirrors to positions very near to predetermined positions. There is no significant difference in the severity of the problem

because the motion is applied to the mirrors rather than the beamsplitter as we are assuming the mirrors to be just large enough so that the beamsplitter will be the aperture stop of the system.

If fabrication were the only problem, then it would be possible to attain more acceptance in an interferometer by use of a larger A as it is presently possible to fabricate larger beamsplitters than diffraction gratings, but it is perhaps more reasonable to assume limitation of installation space, in which case one should compare equal area beamsplitters and gratings. It is also to be noted that diffraction is not as severe a problem in interferometers, so that if there is adequate radiation, one may obtain spectra using smaller instruments. However, the main argument is that $\Omega \cos \alpha$ can be much larger in the interferometer.

Although over narrow wavelength regions, both instruments may have Ω increased by the use of dispersive materials, and spectrometers may have Ω increased by use of a curved slit both of these methods reduce the ability to cover wide spectral regions and therefore are not considered.

To find the tolerable angular slit width, for a spectrometer, it is assumed that the grating is used so that the diffracted radiation is returned to a detector on the same side of the normal as the entrance slit. Then the grating equation is:

$$i\lambda = d(\sin\alpha + \sin\gamma) \qquad \text{(S9-8)}$$

where

i = order number

d = line spacing of grating

α = angle of incidence

γ = angle of diffraction.

By differentiation one may arrive at

$$\Delta\alpha \cos\alpha \approx \left(\frac{\Delta\lambda}{\lambda}\right)_\alpha (\sin\alpha + \sin\gamma) \leq 2\left(\frac{\Delta\lambda}{\lambda}\right)_\alpha \qquad \text{(S9-9)}$$

where

$\Delta\alpha$ = angular slit width

$\left(\frac{\Delta\lambda}{\lambda}\right)_\alpha$ = resolution degradation from use of finite slit width.

The above standard derivation assumes that the direction of incidence lies in a plane normal to the grating and to its rulings. If this is not the case, the equation may be generalized to

$$i\lambda = d(\sin\alpha + \sin\gamma)\cos\beta \qquad \text{(S9-10)}$$

where

β = angle between direction of incidence (and hence diffraction) and above mentioned plane

α = angle between projection of incidence direction into above mentioned plane and grating normal.

Then, assuming β small, and using a series expansion of $\cos\beta$ about $\beta = 0$, one may arrive at

$$\beta = \pm\sqrt{\left(\frac{2\delta\lambda}{\lambda}\right)_\beta} \qquad \text{(S9-11)}$$

The plus or minus sign corresponds to the fact that the angular slit length may extend to either side with identical results. The loss of spectral resolution from slit length (for a given $\frac{\delta\lambda}{\lambda}$) may not be as serious as that from slit width, as it amounts to convolving the instrumental profile with a pointed rather than a flat pulse. However, taking the resolution losses due to finite slit length and finite slit widths as being equal, one finds for the spectrophotometer

$$\Omega \cos\alpha \approx \delta\alpha \cos\alpha \; 2\delta\beta \leq 4\sqrt{2}\left(\frac{\Delta\lambda}{\lambda}\right)^{3/2} \tag{S9-12}$$

For an interferometer, one obtains identical fringes for all wavelengths-angle combinations for which $\cos\theta$ is a constant, where θ is the angle between the axis and direction of propagation. Thus it may be shown that the angular aperture of the interferometer, in both directions, is the same as the angular slit length of the spectrometer. That is

$$\Omega_I \cos\alpha \approx 4\sqrt{2}\,\frac{\Delta\lambda}{\lambda} \tag{S9-13}$$

for $\alpha = 45°$, the usual layout of a Michelson interferometer. By comparison of equations (S9-12)and (S9-13), the interferometer enjoys an advantage of the square root of the spectral resolution in throughput.

Conversion of this advantage into a definite advantage in signal-to-noise ratio or stabilization requirements requires knowledge of the characteristics of the source, and of the detector used. There are two counteracting influences with respect to the detector. First, since detector noise tends to increase with the square root of area[4], an attempt to increase signal-to-noise ratio by use of FTS would not yield an advantage of the order of the square root of the resolution but only the fourth root, for the same focal length in the condenser. However, stigmatic imaging is not required of this element in an interferometer, and hence there is hope of using a small detector with a very short focal length condenser. Investigation of the implications of this trade-off is considered to be beyond the scope of this appendix. However, disregarding short focal length, the following possibilities are available.

1) Keep the same detector area and solid acceptance angle, but reshape it into a square or circle from a rectangle. This would decrease signal-to-noise ratio slightly, cause resolution to be nearer the theoretical maximum and:

a) for a point source, reduce pointing accuracy requirements, **or**

b) for an extended source, (e. g., the sun) cause the spectrum to be averaged over a spatial region with smaller maximum dimension.

2) For an extended source, increase aperture and detector area to increase signal-to-noise ratio. (Note that for a point source, the signal-to-noise ratio is degraded.)

IV. FELLGETT'S ADVANTAGE AND DISTRIBUTION OF NOISE

Section III above discussed how the FTS method provides an instrument capable of accepting more radiation than a grating spectrometer. For very weak sources, FTS also is a more efficient method of using the radiation accepted. The more efficient use is called Fellgett's advantage, and is discussed in this section. The **FTS** sometimes causes a redistribution of noise over wavenumber and this effect is discussed in this section also.

The FTS method is compared in this section to a grating spectrometer to evaluate Fellgett's advantage. The comparison is made for spectra which are essentially flat over the range of interest, e. g., it is to be used to look for weak absorption lines. The effects of detector size and acceptance were discussed previously and in this section attention is focused on the use of the radiation by assuming equal acceptance (perhaps because of use of a point source) and the same detector. It is assumed that the grating delivers the fraction η of the radiation into the order in use, and that the beamsplitter it is compared to has energy transmittance and reflectances of 50%, so that for $\eta = 1$, both instruments would have the same peak output for monochromatic radiation. Also, both instruments are allowed equal time in which to make measurements.

Before proceeding further, it is necessary to develop the concept of the number of degrees of freedom of a spectrum. Assume that a range of wavenumber $\Delta\nu$ is being investigated with a resolution $\delta\nu$ Then the number of degrees of freedom of the spectrum (N) may be defined as

$$N = \frac{\Delta \nu}{\delta \nu}$$

Usually, the range $\Delta \nu$ may be readily defined unambiguously. The resolution will be defined as half the distance from the peak to the first zero of the instrument profile in the case of diffraction limited gratings and FTS with very good collimation and triangular apodization. The N defined above is then the number of independent samples required to specify the spectrum. It is not as easy to define the resolution of the slit width limited spectrometer or of the interferometer accepting light over a wider angle. However, the precise definition of resolution is not critical to the discussion.

Both the spectrometer and the interferometer will require N readings, the spectrometer at N different wavenumbers, and the interferometer at N different separations. Thus, the requirement of the same total time results in the same time per reading.

Because interferometer derived spectra are found by computation, it is necessary to know how the noise propagates through the computations. In the interferometer, the main noise sources will have Poisson or Gaussian fluctuations. For these fluctuations, a simple weighted root sum square method may validly be used to find the root-mean-square fluctuations in a weighted sum.

Letting E be the amount of energy falling on the detector if the grating were 100% efficient, in one time interval and σ_s be the fluctuation in this, it is clear that the signal-to-noise ratio for the spectrometer is

$$\left.\frac{s}{n}\right|_s = \frac{\eta E}{\sigma_s} \qquad \text{(S9-14)}$$

To determine the signal and noise of the interferometer spectrogram, it may be noted that the final spectrogram values are weighted sums of interferometer readings. A wave number interval gives rise to radiant power which may be expressed as a bias plus a cosine function of separation, (see Reference 3). The peak value of this power is the same as that in a

grating spectrometer with $\zeta = 1$, by hypothesis, and the minimum value is zero. Thus the peak power of the cosine function, which is the useful part, is half the peak power of the spectrometer (for $\zeta = 1$) and as the time allotted to measurement is equal to the spectrometer's time, the maximum useful energy in a reading is $\frac{E}{2}$. The resulting spectrogram ordinate is seen to be the weighted sum of N interferogram ordinates of which the largest are $\frac{e}{2}$. The cosine form in the interferometer and the cosine weighting in analysis combine to have an average value of 1/2, and the apodization function has an average value of 1/2. The signal in the interferometer derived spectrogram is

$$s_I = \frac{NE}{8} \tag{S9-15}$$

Similarly, suppose that σ_I is the noise fluctuation in a single interferometer reading. These fluctuations are weighted in the analysis both by cosines, and by apodization which is equivalent to the terms of the sequence $\left\{1, \frac{N-1}{N}, \frac{N-2}{N} \cdots \frac{1}{N}\right\}$. The RMS fluctuation in the **FTS** spectrum is found to be

$$n_I = \sqrt{\frac{N}{6}}\ \sigma_I \tag{S9-16}$$

To finish the comparison, the connection must be made between σ_s and σ_I. There are three cases of interest. The first case considered is that in which the noise is independent of signal. In the second case, the noise is proportional to the square root of the signal, and in the third, the noise is proportional to the signal.

In the first case, the noise is supposed independent of signal, so that σ_I will be the same as σ_s. By using this, in combination with equations (S9-14, 15 and 16), it may be seen that the signal-to-noise ratio of the **FTS** spectrogram is

$$\left.\frac{s}{n}\right|_I = \sqrt{\frac{N}{2}}\left[\frac{\sqrt{3}}{4\xi}\right]\left.\frac{s}{n}\right|_s \qquad \text{(S9-17)}$$

Because nearly perfect beamsplitters are available and $\xi \approx \frac{\sqrt{3}}{4}$ is a reasonable value for ξ, Fellgett's advantage of $\sqrt{\frac{N}{2}}$ * applies in full. Noise independent of signal level is characteristic of certain infrared detectors and of phototubes operated at such low levels that the signal current is small compared to the dark current.

The second case to consider is that in which the noise is proportional to the square root of the signal; this is characteristic of photon counting processes. For the assumed flat spectrum, the radiation level in the spectrometer will be η/N times that of a perfectly blazed grating with monochromatic light, and the level in the interferometer will be half. Thus the level in the interferometer is $N/2\eta$ times that in the spectrometer. It may be seen that the advantage is more than cancelled, the FTS method having only about 65% of the signal-to-noise of the spectrometer.

Finally, if the fluctuations are proportional to the signal level, characteristic of scintillation, the FTS method is at a great disadvantage. However, scintillation noise can be reduced by separating out part of the beam and using it as a reference level.

It should be noted that FTS causes the noise to be redistributed so that it is uniform across the spectrum. Thus it cleans up emission lines and degrades absorption lines if the noise is proportional to the square root of the radiation level. (For noise independent of level, the noise is already uniform, and for noise proportional to level FTS would not be used.)

---

* The factor $\sqrt{\frac{N}{2}}$ is based on a simpler analysis applicable to a few isolated emission lines, assuming $\eta = 1$.

V. CHARACTERISTICS OF INTERFEROMETER SUITABLE FOR FTS IN SPACE

Interferometers for astronomical FTS on earth have usually been quite complicated, with multiple optical paths. These multiple paths have been required to balance out skylight, to provide for correction of scintillations caused by the atmosphere, and to permit calibration paths. In space, skylight will be absent, and since any interferometer will be expected to be usable over a wide spectral range, all optical elements with focusing properties will be reflective. That is, lenses will be replaced by curved mirrors. Hence, the calibration paths can be those shadowed by secondary mirrors. Also, as fluctuations in radiation are expected to be at lower rates caused only by instability in source and vehicle, provision for calibrating them need not be as central to design. Hence, it is concluded that an ordinary Michelson interferometer layout would be adequate for U. V., Visible, and near IR.

Mertz[5] suggests using a second output beam to provide calibration of scintillation. While this may be a useful suggestion in general, it is not particularly applicable to interferometers to be used in space because:

1) A very special type of beamsplitter is required such that certain ratios of reflectances and transmittances are constant over wave number so that normalization can be performed.

2) This process greatly increases the complexity and hence weight and bulk of the interferometer.

A plan for scintillation normalization which is less dependent on the detailed properties of the beamsplitter is to use a symmetrical beamsplitter and obtain calibration by covering one of the mirrors.

The most serious problems of FTS from a space vehicle are mechanical ones. The scanning mirror must be accurately moved through very small distances for spectrograms with many degrees of freedom. They are the only ones for which the advantages of FTS are appreciable. These mechanical problems are aggravated by the requirements that the mechanism must operate under zero apparent gravity and withstand the shock of launch.

The lack of apparent gravity prevents its use to hold carriages in proper relation to ways and the shock of launch requires strength in the method chosen to replace gravity. These space requirements are expected to dictate mechanisms with considerably more friction than those used on terrestrial interferometers. The best results to date have been obtained by P. and J. Connes[6], using an interferometer with a drive which is essentially a linear D. C. electric motor. It is not certain that such a drive can generate adequate force with reasonable current draw in a space interferometer. Consideration should be given to hydraulic drives, both of the type using servomotors and mechanical pumps, and of the type with electromagnetic pumps.

A problem common to spectrometer and interferometer design is that it would be desirable to operate with an evacuated instrument so that ultraviolet spectra could be measured. This problem is expected to be equally serious for interferometers and spectrometers.

Another problem of interferometry in space is the calibration of scanning displacements. Monitoring fringes from a monochromatic standard source in an interferometer rigidly coupled to or in the scanning interferometer is one of the methods used, and possibly the best. Any source used must be capable of withstanding the shock of launch and operation over long periods without gravity or maintenance. As some lasers provide very nearly monochromatic light, and it is sometimes claimed that deterioration in their mirrors is due to the corrosive effects of the atmosphere it would be desirable to life test laser mirrors in high vacuum. It should also be determined if lasers could be sufficiently ruggedized to withstand shock of launch without losing adjustment.

There are other types of interferometers which should be mentioned. The folded Jamin interferometer[5] uses dispersive elements and so may be expected to have a shorter spectral range, also it appears to be heavier than an equivalent Michelson interferometer, however, the motion required is angular rather than linear and this may be expected to ease mechanical and calibration problems. A worthwhile project could be to invent an interfero-meter with the advantages of the folded Jamin, without the requirements of

the large block of transparent material. Also, for far infrared, interferometers using wire grating polarizing beamsplitters are advantageous. These interferometers[5] permit considerable simplification of layout.

VI. CONCLUSIONS

Fourier transform spectrometry offers the following potential advantages over convention spectrometry for use in space.

1) For weak sources, an improved signal-to-noise ratio,

2) For strong extended sources, an improved signal-to-noise ratio or results more applicable to testing current theories,

3) For strong point sources, relaxed stabilization requirements.

However, considerable theoretical and experimental work is required to determine if these potential advantages are attainable and, if so, to secure them.

VII. RECOMMENDATIONS

The following recommendations, in some cases, point out alternate paths to the goal of building interferometers which can be useful in space. Hence, if time is not critical, they may be tried sequentially.

1) Study the characteristics of possible beamsplitter design to find beamsplitters usable over wide wavelength ranges (theoretical).

2) Design, fabricate and test (in simulated space environment) wide range interferometers of the Michelson type. Consideration should be given to both electromechanical-hydraulic and electromagnetic-hydraulic drives (theoretical and experimental).

3) Give further study to the detector size trade-off mentioned above (theoretical).

4) Perform stress analysis to determine if lasers suitable for calibration can withstand launch (theoretical).

5) Life test laser mirrors in high vacuum (experimental).

6) Evaluate the folded Jamin interferometer for space use (theoretical).

7) Attempt to invert other suitable angular interferometers.

VIII. REFERENCES

1. Sawyer, R. A. , Experimental Spectroscopy, New York, Prentice-Hall, 1951.

2. Born, M. and Wolf, E., Principles of Optics, New York, Pergamon Press, Chap. VII and VIII, 1964.

3. Blackman, R. B. and Tukey, J. W., The Measurement of Power Spectra, New York, Dover, 1958.

4. Jones, R. C., "Phenomenological Description of the Response and Detecting Ability of Radiation Detectors", Proc. IRE, Vol. 47, No. 9, p. 1495 FF, September 1959.

5. Mertz, L., Transformations in Optics, New York, Wiley & Sons, Inc., Chap. 1 and 2, 1965.

6. Connes, J. & P., "Near Infrared Spectra by Fourier Spectroscopy, I. Instruments and Results", Journal of the Opt. Soc. of Am., Vol. 50, No. 7, pp. 896-910, July 1966.

APPENDIX B-1

BIBLIOGRAPHY

Abell, G. : Exploration of the Universe. Holt, Rinehart and Winston, New York, 1964.

Abetti, G. (J.B. Sidgwick, trans.): The Sun. The MacMillan Company, New York, 1957.

Alfvén, H. : Cosmical Electrodynamics. Oxford at the Clarendon Press, 1950.

Allen, C. W. : Astrophysical Quantities. Athlone, London, 1955.

Aller, L. H. : The Abundance of the Elements. Interscience Publishers, Inc., New York, 1961.

Aller, L. H. : Astrophysics. Vol. I - The Atmospheres of the Sun and Stars. Second ed. , Ronald Press Company, 1963.

American Optical Company: Feasibility Study of a 120-inch Orbiting Astronomical Telescope. Final Report. J. W. Fecker Div. , Pittsburgh, 1963.

Athay, R. G.; and Thomas, R. N. : Physics of the Solar Chromosphere Interscience Publishers, Inc. , New York, 1961.

Baker, R. A. : Astronomy. 8th ed., D. Van Nostrand, 1964.

Barnes, C. W. : Object Restoration in a Diffraction-Limited Imaging System. J. Opt. Soc. Am. , vol. 56, no. 5, May 1966, pp 575ff.

Baum, W. A. : The Detection and Measurement of Faint Astronomical Sources. Astronomical Techniques, W. Hiltner, ed. Vol. II of Stars and Stellar Systems, G. P. Kuiper and B. M. Middlehurst, ed. , University of Chicago Press, 1962.

Beer, A. , ed. : Vistas in Astronomy. Vol. I and 11. Pergamon Press, 1956.

Berkner, L. V. ; and Odishaw, Hugh, ed. : Science in Space. McGraw-Hill Book Co. , Inc., New York, 1961.

Blackman, R. B. ; and Tukey, J. W. : The Measurement of Power Spectra. Dover, New York, 1958.

Blanco, V. M. ; and McCuskey, S. W. : Basic Physics of the Solar System. Addison-Wesley Pub. Co., Massachusetts, 1961.

Born, M.; and Wolf, E.: Principles of Optics. Second ed., MacMillan Company, New York, 1964.

Bracewell, R. N., ed.: Paris Symposium on Radio Astronomy, 1958. Stanford University Press, 1959.

Bray, R. J.; and Loughhead, R. E.: Sunspots. John Wiley & Sons, Inc., New York, 1965.

Brück, H. A.; and Brück, M. F.: The Solar Installation of Dunsink Observatory. Vol. I of Vistas in Astronomy, sec. 5, A. Beer, ed., Pergamon Press, London, 1955.

Buchheim, R. W.; and the Staff of the Rand Corporation: New Space Handbook: Astronautics and Its Applications. Vintage Books, New York, 1963.

Campbell, C. E.: Project TECH TOP - Study of Lunar, Planetary and Solar Topography. First Phase Report. Cornell Aeronautical Laboratory, Inc. Rept. VC-2104-D-1, August 1966. (NASA CR-650).

Chandrasekhar, S.: An Introduction to the Study of Stellar Structure. The University of Chicago Press, Chicago, 1939.

Chandrasekhar, S.: Plasma Physics. The University of Chicago Press, Chicago, 1960.

Chandrasekhar, S.: Radiative Transfer. Oxford at the Clarendon Press, 1950.

Chandrasekhar, S.; and Breen, F.: Ap. J., vol. 104, 1946, p. 430.

Chaney, L. W.; and Loh, L. T.: An Infrared Interference Spectrometer - Its Evaluation, Test, and Calibration. NASA CR-61, June 1964.

Condon, E. V.; and Odishaw, H., ed.: Handbook of Physics. McGraw-Hill Book Company, 1958.

Connes, J. and P.: Near-Infrared Spectra by Fourier Spectroscopy. I. Instruments and Results. J. Opt. Soc. Am., vol. 56, no. 7, July 1966, pp. 896-910.

Conrady, A. E.: Applied Optics and Optical Design. Dover, New York, 1957.

Corliss, W. R.: Space Probes and Planetary Exploration. D. Van Nostrand Company, Inc., Princeton, 1965.

Cotton, A.: Ap. J., vol. 33, 1911, p. 375.

Cowling, T. G.: Magnetohydrodynamics. Interscience, New York, 1957.

Cowling, T. G. : Solar Electrodynamics. The Solar System. Vol. I - The Sun, ch. 8, G. P. Kuiper, ed., University of Chicago Press, 1953.

Cutrona, L. J. ; et al.: Optical Data-Processing and Filtering Systems. IRE Trans. on Inf. Theory IT-6, June 1960.

Dickstein, D. H. : A Solar Probe. IAS Paper 61-179-1873.

Ditchburn, R. W. : Light. Second ed., Interscience Publishers, Inc., New York, 1963.

Dugan, D. W. : A Preliminary Study of a Solar Probe Mission. Ames Research Center. NASA Technical Note D-783, April 1961.

Eberhard, G. ; Kohlschutter, A. ; and Ludendorff, H., ed. : Handbuch der Astrophysik. 4 vols., Springer, Berlin, 1929.

Eckman, D. P., ed. : Systems: Research and Design - Proceedings of the First Systems Symposium at Case Institute of Technology. John Wiley & Sons, Inc., New York, 1961.

Eddington, A. S. : The Internal Constitution of the Stars. Cambridge, 1926.

Edelbaum, T. N. : Some Extensions of the Hohmann Transfer Maneuver. ARS Journal, vol. 29, no. 11, November 1959.

Edmonds, F. N. : Ap. J., vol. 131, 1960.

Evans, A. E.; et al.: Solar Panel Design Considerations. ARS Space Power Systems, vol. 4, 1960.

Ferraro, V. C. A. : M. N., vol. 97, 1937, p. 458.

Flügge, S., ed. : Handbuch der Physik. Vol. L-LIV, Astrophysik, Springer, Berlin, 1959.

Gamow, G. : A Star Called the Sun. The Viking Press, Inc., New York, 1964.

Garwin, R. L. : Solar Sailing - A Practical Method of Propulsion Within the Solar System. ARS Jet Propulsion, vol. 28, no. 3, 1958.

Glasstone, S. : Sourcebook on the Space Sciences. D. Van Nostrand Co., Inc., Princeton, 1965.

Goldberg, L. ; and Aller, L. A. : Atoms, Stars and Nebulae. Blakiston, Philadelphia, 1943.

Goldstein, B. S. : Radar Range Equations for Narrow-band Optical Sources. MIT, Linc. Lab., Tech. rept. 319, June 14, 1963.

Gorokhovskii, L. N. (Grace E. Lockie, trans. from Russian): Spectral Studies of the Photographic Process. Focal Press, New York, 1960.

Greenstein, J. L., ed. : Stellar Atmospheres. Vol. VI of Stars and Stellar Systems, Kuiper & Middlehurst, ed., Chicago, 1960.

Hall, C. F. ; Nothwang, G. J. ; and Hornby, H. : A Feasibility Study of Solar Probes. Institute of the Aerospace Sciences Paper no. 62-21.

Hardy, A. ; and Perrin, F. : The Principles of Optics. McGraw-Hill Book Company, New York, 1932.

Harris, J. S., Sr. : Image Evaluation and Restoration. J. Opt. Soc. Am., vol. 46, no. 5, may 1966, p. 569ff.

Helstrom, C. W. : Statistical Theory of Signal Detection. Pergamon Press, New York, 1960.

Herzberg, G. : Atomic Spectra and Atomic Structure. Dover Publications, New York, 1944.

Holladay, T. H. ; and Gallatin, J. D. : Phase Control by Polarization in Coherent Spatial Filtering. J. Opt. Soc. Am., vol. 56, no. 7, , July 1966, p. 869,

Hopkins, H. H. : The Frequency Response of a Defocused Optical System. Proc. Roy. Soc. (London), Series A, vol. 231, 1955, p. 91ff.

Hoyle, F. : Recent Researches in Solar Physics. Cambridge, London, 1949.

Hulst, Hendrick van de: Light Scattering by Small Particles. John Wiley & Sons, New York, 1957.

Hynek, J. A. : Astrophysics. McGraw-Hill, New York, 1951.

Jenkins, F. A. ; and White, H. E. : Fundamentals of Optics. McGraw-Hill Book Company, New York, 1957.

Jones, R. C. : Phenomenological Description of the Response and Detecting Ability of Radiation Detectors. Proc. IRE., vol. 47, no. 9, September 1959, p. 1495ff.

Kallmann, H., ed. : Space Research. North Holland, Amsterdam, 1960.

Kawabata, K. : Publ. Astron. Soc., Japan, vol. 12, 1960.

Kinzly, R. E. ; Roetling, P. G. ; and Holladay, T. M. : Project SLOPE - Study of Lunar Orbiter Photographic Evaluation. Phase I Final Report. Cornell Aeronautical Laboratory, Inc. , Rept. VS-2182-D-2, May 20, 1966. (NASA CR-66158).

Kourganoff, V. (in collaboration with I. W. Busbridge): Basic Methods in Transfer Problems: Radiative Equilibrium and Neutron Diffusion. Oxford at the Clarendon Press, 1952.

Kovasznay, L. S. G. ; and Joseph, H. M. : Image Processing. Proc. IRE, vol. 43, May 1955, pp. 560-570.

Kuiper, G. P., ed. : The Atmospheres of the Earth and Planets. The University of Chicago Press, 1949.

Kuiper, G. P. ; and Middlehurst, B. M. : Telescopes. University of Chicago Press, 1960.

Lansraux, G. : Diffraction Instrumental. Editions de la Revue d'Optique, Paris, 1953.

LeGalley, D. P. ; and McKee, J. W. : Space Exploration. McGraw-Hill Book Company, New York, 1964.

Lehnert, B., ed. : Electromagnetic Phenomena in Cosmical Physics. Cambridge, 1958.

Leighton, R. B. ; Noyes, R. W. ; and Simon, G. W. : Ap. J., vol. 135, 1962.

Levin, M. J. : Optimum Estimation of Impulse Response in the Presence of Noise. IRE Convention Record, pt. 4, 1959, pp. 174-181.

Liller, W., ed. : Space Astronomy. McGraw-Hill Book Company, New York, 1961.

Lyot, B. : Ann d'Astrophys., vol. 7, 1944.

Macris, C. J. ; and Banos, G. J. : Mean Distant Between Photospheric Granules and Its Change with the Solar Activity. Men. Nat. Obs. Atheres: Serial 1, no. 8, 1961.

Massey, Sir Harrie: Space Physics. Cambridge at the University Press, 1964.

McClelland, D. H. : Solar Concentrators for High Temperature Space Power Systems. ARS Space Power Systems, vol. 4, 1960.

McCrea, W. H. : Physics of the Sun and Stars. Hutchinson House, London, 1950.

McMath, R. R. ; and Mohler, O. C. : Solar Instruments. Handbuch der Physik, vol. LIV. Springer-Verlag, Berlin, 1962.

Mees, C. E. K. : Theory of the Photographic Process. MacMillan, New York, 1954, pp. 852-854.

Menzel, D. H. : Our Sun. Second Ed., (Harvard Books on Astronomy), Harvard University Press, 1959.

Menzel, D. H. ; and Pekeris, C. L. : M.N., vol. 96, 1936.

Mertz, L. : Transformations in Optics. Wiley, New York, 1965.

Miczaika, G. R. ; and Sinton, W. M. : Tools of the Astronomer. Harvard University Press, Massachusetts, 1961.

Minnaert, M. : The Photosphere. The Solar System. Vol. I - The Sun, ch. 3, G. P. Kuiper, ed. , University of Chicago Press, 1953.

Moore, C.E. : The Identification of Solar Lines. The Solar System. Vol. I - The Sun, ch. 4, G. P. Kuiper, ed., University of Chicago Press, 1953.

Motz, L., ed. : Astronomy A to Z. Grosset and Dunlap, New York, 1964.

Newton, H. W. ; and Nunn, M. L. : M. N., vol. 111, 1951, p. 413.

Offner, A. : Ray Tracing Through a Holographic System. J. Opt. Soc. Am,, vol. 56, no. 11, November 1966, p. 1509.

O'Neill, E. L. : Introduction to Statistical Optics. Addison-Wesley, Reading, Massachusetts, 1963, p. 163.

O'Neill, E. L. : Spatial Filtering in Optics. IRE Trans. on Inf. Theory, IT-2, 1956, p. 56ff.

Ordway, F. I. : Advances in Space Science and Technology. 4 vol., Academic, New York, 1959-1962.

Pawsey, J. L. ; and Bracewell, R. N. : Radio Astronomy. Oxford at the Clarendon Press, 1955.

Payne-Gaposchkin, C. : Introduction to Astronomy. Prentice-Hall, Inc. , New Jersey, 1954.

Perlis, S. : Theory of Matrices. Addison-Wesley, Reading, Mass., 1952.

Phillips, D. L. : A Technique for the Numerical Solution of Certain Integral Equations of the First Kind. J. Assoc. for Computing Machinery, vol. 9, 1962, pp. 84-97.

Rapport, S. ; and Wright, H. : Astronomy. Washington Square Press, Inc. , New York, 1964.

Richtmyer, F. K. ; and Kennard, E. H. : Introduction to Modern Physics. McGraw-Hill, Inc., New York, 1947.

Roetling, P. G. : Effects of Signal-Dependent Granularity. J. Opt. Soc. Am., vol. 55, no. 1, January 1965, pp. 67-71.

Roetling, P. G. ; and Hammill, H. B. : Study of Spatial Filtering by Optical Diffraction for Pattern Recognition. Final Report. Cornel1 Aeronautical Laboratory, Inc., Rept. no. VE-1522-G-1. 13 February 1962. (RADC-TR-62-93).

Rubashov, B. M. : Problems of Solar Activity. NASA TT F-244, 1964.

Russell, H. N. ; Dugan, R. S. ; and Stewart, J. Q. : Astronomy. 2 vol., Ginn & Co., Boston, 1927.

Sawyer, R. A. : Experimental Spectroscopy. Prentice-Hall, New York, 1951.

Schade, O. H. : Optical Image Evaluation. Nat. Bur. Stand. Circ. 526, 1954, p. 231ff.

Schwarzschild, M. : Structure and Evolution of the Stars. Princeton, 1958.

Schwarzschild, M. : Ap. J., vol. 130, 1959, p. 345.

Seifert, H. : Space Technology. J. Wiley & Sons, New York, 1959.

Shane, C. D. : Bull. Lick Obs., vol. 16, 1932, p. 76.

Shannon, C. E. ; and Weaver, W. : The Mathematical Theory of Communication. The University of Illinois Press, Urbana, Illinois, 1949.

Shapley, Harlow, ed. : Source Book in Astronomy, 1900-1950. Harvard University Press, Cambridge, Mass. , 1960.

Smith, R. A. ; Jones, F. E. ; and Chasmar, R. R. : The Detection and Measurement of Infrared Radiation. Oxford, 1957.

Solodovnikov, V. V. : Introduction to the Statistical Dynamics of Automatic Control Systems. ch. V., Dover, New York, 1960.

Sommerfeld, A. : Optics. Academic Press, New York, 1964.

St. John, C. E. : Ap. J., vol. 67, 1928, p. 195.

Starr, V. P. ; and Gelman, P. A. : Energetics of the Solar Rotation. Ap. J., vol. 141, 1964.

Stratton, F. J. M. : Astronomical Physics. Methuen & Co., Ltd., London, 1925.

Stroke, G. W. : An Introduction to Coherent Optics and Holography. Academic, New York, 1966.

Strömgren, Bengt: Aufgaben und Probleme der Astrophotometrie. Astrophysik. Band 26 of Handbuch der Experimentalphysik, Akademische Verlagsgesellschaft M. B. H., Leipzig, 1937.

Struve, O.: Stellar Evolution. Princeton, 1950.

Sweet, P. A.: M. N., vol. 109, 1949, p. 507.

Swerling, P.: Parameter Estimation Accuracy Formulas. IEEE Transactions on Information Theory, vol. IT-10, no. 4, p. 306.

Thompson, M. M., ed.: Manual of Photogrammetry, American Society of Photogrammetry, Falls Church, Va., 1966.

Trabka, E. A.; and Roetling, P. G.: Image Transformation for Pattern Recognition using Incoherent Illumination and Bipolar Aperture Masks. J. Opt. Soc. Am., vol. 54, no. 10, Oct. 1964, pp. 1241-1252.

Turin, G. L.: An Introduction to Matched Filters. IRE Transactions on Information Theory, vol. IT-6, no. 3, June 1960.

Turin, G. L.: On the Estimation in the Presence of Noise of the Impulse Response of a Random, Linear Filter. IRE Transactions on Information Theory, vol. IT-3, March 1957, pp. 5-10.

Unsöld, A.: Physik der Sternatmospharen. 2nd ed. Springer, Berlin, 1955.

Valley, S. L., ed.: Handbook of Geophysics and Space Environments. Air Force Cambridge Research Laboratories, 1965.

Valley, S. L., ed.: Space and Planetary Environments. AFCRL Rept. No. 139, Jan. 1962.

Vaucouleurs, Gerard de (R. W. Wright, trans.): Astronomical Photography. The MacMillan Co., 1961.

Walen, C.: Ark. f. Mat., Astro. Och. Pys., vol. 33A, no. 18, 1946.

Ward, F.: J. Pure and Appl. Geophys., vol. 58, 1964, p. 61.

Ward. F.: Ap. J., vol. 141, 1965, p. 2.

Wayman, P. A.: Optical Systems Employed in Spectroheliographs. Vol. I of Vistas in Astronomy, sec. 5, A. Beer, ed., Pergamon Press, London, 1955.

Weil, N. A.: Lunar and Planetary Surface Conditions. Academic Press, New York, 1965.

Whitney, C.: Granulation and Oscillations of the Solar Atmosphere. Smithsonian Contributions to Astrophysics, vol. 2, no. 12. Smithsonian Institution, 1958.

Wiener, N. : Extrapolation, Interpolation, and Smoothing of Stationary Time Series. Wiley, New York, 1949.

Wilson, O.C. : Science. vol. 122, Lancaster, Pa., 1955, p. 882.

Wolf, M. : Limitations and Possibilities for Improvement of Photovoltaic Solar Energy Converters. Proc. of IRE, vol. 7, 1960.

Woolley, R. v. d. R. ; and Stibbs, D. W. N. : The Outer Layers of a Star. Oxford at the Clarendon Press, 1953.

Wrubel, M. H. : Stellar Interiors. Handbuch der Physik, vol. LI. Springer-Verlag, Berlin, 1958.

Zirker, J. B. : Ap. J., vol. 135, 1962.

Anon: Manual of Physical Properties of Kodak Aerial and Special Sensitized Materials. Eastman Kodak Co. , Rochester, New York. (Published as a loose-leaf notebook and revised a chapter at a time.)

Anon: Significant Achievements in Space Astronomy, 1958-1964. NASA SP-91, 1966.

Anon: Significant Achievements in Space Communications and Navigation, 1958-1964. NASA SP-93, 1966.

Anon: Significant Achievements in Satellite Geodesy, 1958-1965. NASA SP-94, 1966.

Anon: Significant Achievements in Ionospheres and Radio Physics, 1958-1964. NASA SP-95, 1966.

Anon: Significant Achievements in Satellite Meteorology, 1958-1964. NASA SP-96, 1966.

Anon: Significant Achievements in Solar Physics, 1958-1964. NASA SP-100, 1966.